

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2003-526376

(P2003-526376A)

(43) 公表日 平成15年9月9日 (2003.9.9)

|                             |       |              |            |
|-----------------------------|-------|--------------|------------|
| (51) Int.Cl. <sup>7</sup>   | 識別記号  | F I          | テームト* (参考) |
| C 1 2 N 15/09               | Z N A | C 1 2 N 1/15 | 4 B 0 2 4  |
| 1/15                        |       | 1/19         | 4 B 0 6 5  |
| 1/19                        |       | 1/21         |            |
| 1/21                        |       | 15/00        | Z N A A    |
| 5/10                        |       | 5/00         | B          |
| 審査請求 未請求 予備審査請求 有 (全 108 頁) |       |              |            |

(21) 出願番号 特願2001-567365(P2001-567365)  
 (86) (22) 出願日 平成13年3月13日 (2001.3.13)  
 (85) 翻訳文提出日 平成14年9月13日 (2002.9.13)  
 (86) 国際出願番号 P C T / U S 0 1 / 0 7 8 7 0  
 (87) 国際公開番号 W O 0 1 / 0 6 8 8 8 2  
 (87) 国際公開日 平成13年9月20日 (2001.9.20)  
 (31) 優先権主張番号 0 9 / 5 2 5 , 1 6 0  
 (32) 優先日 平成12年3月14日 (2000.3.14)  
 (33) 優先権主張国 米国 (U S)

(71) 出願人 トランスカーヨティック・セラピーズ・インコーポレーテッド  
 アメリカ合衆国マサチューセッツ州02139,  
 ケンブリッジ, アルバニー・ストリート  
 195  
 (72) 発明者 イヴァノフ, エヴゲニ  
 アメリカ合衆国マサチューセッツ州02067,  
 シャロン, イースト・フォックスボロ・ス  
 トリート 390  
 (74) 代理人 弁理士 社本 一夫 (外5名)

最終頁に続く

(54) 【発明の名称】 相同組換えを改善する方法

(57) 【要約】

本発明は、標的DNA中、例えば細胞の染色体DNA中の選択される部位での改変を促進する方法を特徴とする。該方法は、その部位で：(1) 選択されるDNA配列を含む二本鎖DNA配列；(b) 相同組換えを増進する剤、例えばRad52タンパク質またはその機能する断片；および(c) 非相同端連結を阻害する剤、例えば抗Ku抗体またはKu結合オリゴマーもしくはポリマーなどのKuを不活性化する剤を提供し、そして改変が起こるのを可能にすることを含む。非相同端連結を阻害する剤、例えば抗Ku抗体などのKu不活性化剤は、好ましくは、局所で提供される。構成要素(a)、(b)、および(c)は、好ましくは共に投与してもよいし、または別個に投与してもよい。



**【特許請求の範囲】**

【請求項1】 細胞における標的配列の改変を促進する複合体であって：二本鎖DNA配列、相同組換え増進剤、および非相同端連結を阻害する剤を含む、前記複合体。

【請求項2】 相同組換え増進剤が：Rad52タンパク質または機能するその断片、Rad51タンパク質または機能するその断片、Rad54タンパク質または機能するその断片、およびそれらの組み合わせからなる群より選択される、請求項1の複合体。

【請求項3】 相同組換え増進剤がRad52タンパク質または機能するその断片である、請求項1の複合体。

【請求項4】 非相同端連結を阻害する剤が、hMrcl1を不活性化する剤、hRad50を不活性化する剤、Nbs1を不活性化する剤、hLig4を不活性化する剤、hXrcc4を不活性化する剤、およびKuを不活性化する剤からなる群より選択される、請求項1の複合体。

【請求項5】 非相同端連結を阻害する剤がKu不活性化剤である、請求項1の複合体。

【請求項6】 Ku不活性化剤が：抗Ku抗体、Ku結合オリゴマー、およびKu結合ポリペプチドからなる群より選択される、請求項5の複合体。

【請求項7】 Ku不活性化剤が抗Ku抗体である、請求項5の複合体。

【請求項8】 DNA配列が直鎖DNA配列を含む、請求項1の複合体。

【請求項9】 DNA配列に少なくとも1つのターゲティング配列が隣接する、請求項1の複合体。

【請求項10】 DNA配列が外因性制御配列を含む、請求項1の複合体。

【請求項11】 制御配列がプロモーター、エンハンサー、上流活性化配列、足場付着領域または転写因子結合部位である、請求項10の複合体。

【請求項12】 制御配列がプロモーターおよびエンハンサーである、請求項11の複合体。

【請求項13】 制御配列がプロモーターおよび上流活性化配列である、



請求項11の複合体。

【請求項14】 Rad52タンパク質または機能するその断片が、DNA配列上にコーティングされている、請求項3の複合体。

【請求項15】 Rad52タンパク質またはその断片がヒトRad52である、請求項3の複合体。

【請求項16】 抗Ku抗体が抗Ku70抗体である、請求項7の複合体。

【請求項17】 抗Ku抗体が抗Ku80抗体である、請求項7の複合体。

【請求項18】 少なくとも1つの抗Ku抗体が、DNA配列に共有結合している、請求項7の複合体。

【請求項19】 少なくとも1つの抗Ku抗体が、相同組換えを増進する剤に共有結合している、請求項7の複合体。

【請求項20】 複合体が抗Ku70抗体および抗Ku80抗体を含む、請求項7の複合体。

【請求項21】 DNA配列が制御配列を含み、そしてターゲッティング配列がFISH $\beta$ 遺伝子タンパク質コード領域の5'の配列に由来する、請求項9の複合体。

【請求項22】 DNA配列が制御配列を含み、そしてターゲッティング配列がIFN $\alpha$ 遺伝子タンパク質コード領域の5'の配列に由来する、請求項9の複合体。

【請求項23】 DNA配列が制御配列を含み、そしてターゲッティング配列がGCSF遺伝子タンパク質コード領域の5'の配列に由来する、請求項9の複合体。

【請求項24】 ミスマッチ修復タンパク質を阻害する剤をさらに含む、請求項1の複合体。

【請求項25】 細胞の標的DNAにおいて、選択される部位での改変を促進する方法であって：

細胞に、二本鎖DNA配列、相同組換えを増進する剤、および非相同端連結を



阻害する剤を導入し、それによって、染色体DNAの改変を促進し、それによって、染色体DNAにおいて、選択される部位での改変を促進することを含む、前記方法。

【請求項26】 DNA配列が直鎖DNA配列を含む、請求項25の方法。

【請求項27】 DNA配列に少なくとも1つのターゲッティング配列が隣接する、請求項25の方法。

【請求項28】 DNA配列が外因性制御配列を含む、請求項25の方法。

【請求項29】 制御配列が：プロモーター、エンハンサー、上流活性化配列、足場付着領域および転写因子結合部位からなる群より選択される、請求項28の方法。

【請求項30】 制御配列がプロモーターおよびエンハンサーである、請求項28の方法。

【請求項31】 制御配列がプロモーターおよび上流活性化配列である、請求項28の方法。

【請求項32】 相同組換えを増進する剤が：Rad52タンパク質または機能するその断片、Rad51タンパク質または機能するその断片、Rad54タンパク質または機能するその断片、およびそれらの組み合わせからなる群より選択される、請求項25の方法。

【請求項33】 相同組換えを増進する剤がRad52タンパク質または機能するその断片である、請求項25の方法。

【請求項34】 Rad52タンパク質または機能するその断片が、DNA配列上にコーティングされている、請求項33の方法。

【請求項35】 Rad52タンパク質または機能するその断片がヒトRad52である、請求項33の方法。

【請求項36】 非相同端連結を阻害する剤が、hMre11を不活性化する剤、hRad50を不活性化する剤、Nbs1を不活性化する剤、hLig4を不活性化する剤、hXrcc4を不活性化する剤、およびKuを不活性化す



る剤からなる群より選択される、請求項25の方法。

【請求項37】 非相同端連結を阻害する剤がK<sub>u</sub>不活性化剤である、請求項25の方法。

【請求項38】 K<sub>u</sub>を不活性化する剤が、抗K<sub>u</sub>抗体、K<sub>u</sub>結合オリゴマー、およびK<sub>u</sub>結合ポリペプチドである、請求項37の方法。

【請求項39】 K<sub>u</sub>を不活性化する剤がK<sub>u</sub>アンチセンス分子である、請求項37の方法。

【請求項40】 K<sub>u</sub>を不活性化する剤が抗K<sub>u</sub>抗体である、請求項37の方法。

【請求項41】 抗K<sub>u</sub>抗体が抗K<sub>u</sub>70抗体である、請求項40の方法。

【請求項42】 抗K<sub>u</sub>抗体が抗K<sub>u</sub>80抗体である、請求項40の方法。

【請求項43】 少なくとも1つの抗K<sub>u</sub>抗体が、DNA配列に共有結合している、請求項40の方法。

【請求項44】 少なくとも1つの抗K<sub>u</sub>抗体が、Rad52タンパク質またはその断片に共有結合している、請求項40の方法。

【請求項45】 細胞が真菌、植物または動物起源のものである、請求項25の方法。

【請求項46】 細胞が脊椎動物起源のものである、請求項45の方法。

【請求項47】 細胞が初代または二次哺乳動物細胞である、請求項46の方法。

【請求項48】 細胞が初代または二次ヒト細胞である、請求項46の方法。

【請求項49】 細胞が不死化哺乳動物細胞である、請求項46の方法。

【請求項50】 細胞が不死化ヒト細胞である、請求項46の方法。

【請求項51】 DNA配列、相同組換えを増進する剤および非相同端連結を阻害する剤を、複合体として細胞に導入する、請求項25の方法。

【請求項52】 ミスマッチ修復タンパク質を阻害する剤を導入すること



をさらに含む、請求項25の方法。

【請求項53】 ミスマッチ修復タンパク質が：Msh2、Msh6、Msh3、Mlh1およびPMS2からなる群より選択される、請求項52の方法。

【請求項54】 剤がミスマッチ修復タンパク質の発現を阻害する剤である、請求項52の方法。

【請求項55】 剤が抗ミスマッチ修復タンパク質抗体である、請求項54の方法。

【請求項56】 少なくとも1つの抗ミスマッチ修復タンパク質抗体が、DNA配列に共有結合されている、請求項54の方法。

【請求項57】 少なくとも1つの抗ミスマッチ修復タンパク質抗体が、Rad52タンパク質またはその断片に共有結合されている、請求項55の方法。

【請求項58】 DNA配列が制御配列を含み、そしてターゲッティング配列がFSH $\beta$ コード領域の5'の領域に由来する、請求項27の方法。

【請求項59】 DNA配列が制御配列を含み、そしてターゲッティング配列がIFN $\alpha$ コード領域の5'の領域に由来する、請求項27の方法。

【請求項60】 DNA配列が制御配列を含み、そしてターゲッティング配列がGCSFコード領域の5'の領域に由来する、請求項27の方法。

【請求項61】 標的DNAが、野生型配列と異なる10未満の塩基対を有する突然変異を含む、請求項25の方法。

【請求項62】 突然変異が点突然変異である、請求項61の方法。

【請求項63】 DNA配列が、突然変異を訂正することが可能な野生型配列を含む、請求項62の方法。

【請求項64】 標的DNAが嚢胞性線維症膜貫通制御因子（CFTR）遺伝子である、請求項63の方法。

【請求項65】 突然変異が、CFTR遺伝子コード領域のコドン508にコードされるアミノ酸を変化させる、請求項64の方法。

【請求項66】 標的DNAが $\beta$ -グロビン遺伝子である、請求項63の



方法。

【請求項67】 突然変異が、 $\beta$ -グロビン遺伝子コード領域のコドン6にコードされるアミノ酸を変化させる、請求項66の方法。

【請求項68】 標的DNAが因子V I I I 遺伝子である、請求項63の方法。

【請求項69】 突然変異が、因子V I I I 遺伝子コード領域のコドン2209にコードされるアミノ酸を変化させる、請求項68の方法。

【請求項70】 突然変異が、因子V I I I 遺伝子コード領域のコドン2229にコードされるアミノ酸を変化させる、請求項68の方法。

【請求項71】 標的DNAが因子I X 遺伝子である、請求項63の方法。

【請求項72】 標的DNAがフォンウィルブランド因子遺伝子である、請求項63の方法。

【請求項73】 標的DNAが色素性乾皮症グループG 遺伝子である、請求項63の方法。

【請求項74】 請求項25の方法によって作成される相同組換え細胞。

【請求項75】 細胞において、遺伝子のタンパク質コード配列の発現を改変する方法であって：

細胞に、DNA配列が制御配列を含む請求項1の複合体を導入し；

標的とされるゲノム配列の改変を可能にして相同組換え細胞を産生する条件下で、細胞を維持し；そして

制御配列の調節下、遺伝子のタンパク質コード配列の発現を可能にして、それにより遺伝子のタンパク質コード配列の発現を改変する条件下で、相同組換え細胞を維持する

ことを含む、前記方法。



## 【発明の詳細な説明】

## 【0001】

発明の背景

療法タンパク質を投与することによって疾患を治療する、現在のアプローチには、慣用的な薬剤搬送（例えば静脈内、皮下、または筋肉内注射）のための療法タンパク質の *in vitro* 産生が、そしてより最近では遺伝子治療が含まれる。

## 【0002】

療法目的のタンパク質は、療法目的のタンパク質をコードする外因性DNAを、適切な細胞に導入することによって、産生可能である。例えば、療法タンパク質をコードする外因性DNAを含むベクターを細胞に導入し、そしてコードされるタンパク質を発現させることが可能である。ジーンターゲッティングによって、内因性細胞遺伝子およびその発現が修飾可能であることもまた、示唆されてきている。例えば米国特許第5,272,071号、米国特許第5,641,670号、WO 91/06666、WO 91/06667およびWO 90/11354を参照されたい。

発明の概要

本発明は、部分的に、標的とされる部位で、DNA配列に十分にごく近接して、相同組換えを増進する剤、例えばRad52、および非相同端連結を阻害する剤、例えばKu不活性化剤を提供することによって促進される、二本鎖DNA配列および選択される標的DNA、例えば細胞中の染色体DNA間の相同組換えの使用に基づく。Rad52およびKu不活性化剤両方の存在下では、非存在下より、より高い率の相同組換えが起こると予測される。さらに、選択されるDNA配列をテンプレートとして用いて、DNA中の標的とされる部位、例えば細胞中の染色体DNA中の標的とされる部位を改変することを目的とするジーンターゲッティングは、Rad52タンパク質およびKu不活性化剤、例えば抗Ku抗体を提供することによって、促進可能であると予測される。Rad52タンパク質およびKu不活性化剤を、選択されるDNA配列および標的部位にごく近接して提供することによって、Rad52タンパク質およびKu不活性化剤、例えば抗Ku抗体の非存在下よりも、ジーンターゲッティングによる、より高い率の改変



が生じる。

### 【0003】

したがって、1つの側面において、本発明は、標的DNA、例えば細胞の染色体DNAにおいて選択される部位の改変を促進する方法を特徴とする。該方法は、その部位で：(a) 選択されるDNA配列を含む二本鎖DNA配列；(b) 相同組換えを増進する剤、例えばRad52タンパク質または機能するその断片、あるいはRad52または機能するその断片をコードするDNA配列；および(c) 非相同端連結を阻害する剤、例えばKuを不活性化する剤を提供し、そして改変が起こるのを可能にすることを含む。好ましい態様において、部位の改変、例えば選択されるDNA配列および標的DNA間の相同組換えまたは遺伝子訂正が、供給される相同組換え増進剤および非相同端連結阻害剤の非存在下で起こるであろう率より高い率で起こるために、相同組換えを増進する剤の濃度および非相同端連結を阻害する剤の濃度が、選択されるDNA配列および標的DNA間の相互作用部位で、十分であるように、構成要素(a)、(b)、および(c)が提供され、例えば細胞に導入される。非相同端連結を阻害する剤は、好ましくは、局所に提供される。好ましくは、非相同端連結を阻害する剤は、抗Ku抗体などのKu不活性化剤である。

### 【0004】

構成要素(a)、(b)、および(c)は、好ましくは共に導入してもよいし、または別個に導入してもよい。さらに、構成要素の2つを共に導入し、そして第三の構成要素を別個に導入してもよい。例えば、DNA配列および相同組換えを増進する剤、例えばRad52を共に導入してもよいし、またはDNA配列および非相同端連結を阻害する剤、例えばKu不活性化剤を共に導入してもよい。別の好ましい態様において、相同組換えを増進する剤および非相同端連結を阻害する剤を共に導入してもよい。

### 【0005】

構成要素の2つ、または好ましくはすべてを、複合体として提供してもよい。好ましい態様において、該方法は、標的DNAと：(a) 選択されるDNA配列を含む二本鎖DNA配列；(b) 相同組換えを増進する剤、例えばRad52タ



ンパク質または機能するその断片；および（c）非相同端連結を阻害する剤、例えば抗Ku抗体またはKu結合オリゴマーもしくはポリマーなどのKu不活性化剤を含む複合体とを、例えば該複合体を細胞内に導入することによって接触させることを含む。

#### 【0006】

好ましい態様において、1つ、またはそれ以上、好ましくはすべての構成要素は、局所搬送、例えばマイクロインジェクションによって提供され、そして標的ゲノムまたは別の核酸から発現されない。特に好ましい態様において、非相同端連結を阻害する剤、例えばKu阻害剤は、局所搬送、例えばマイクロインジェクションによって提供され、そして標的ゲノムまたは別の核酸から発現されない。

#### 【0007】

好ましい態様において、非相同端連結を阻害する剤は：hMr e 1 1を不活性化する剤、例えば抗hMr e 1 1抗体またはhMr e 1 1結合オリゴマーもしくはポリマー；hR a d 5 0を不活性化する剤、例えば抗hR a d 5 0抗体またはhR a d 5 0結合オリゴマーもしくはポリマー；Nb s 1を不活性化する剤、例えば抗Nb s 1抗体またはhNb s 1結合オリゴマーもしくはポリマー；ヒトリガーゼ4（hL i g 4）を不活性化する剤、例えば抗hL i g 4抗体またはhL i g 4結合オリゴマーもしくはポリマー；hX r c c 4を不活性化する剤、例えば抗hX r c c 4抗体またはhX r c c 4結合オリゴマーもしくはポリマー；Rap 1のヒト相同体（h o m o l o g）を不活性化する剤、例えばRap 1のヒト相同体に対する抗体またはRap 1のヒト相同体に結合するオリゴマーもしくはポリマー；S i r 2 3 0 4のヒト相同体を不活性化する剤、例えばS i r 2 3 0 4のヒト相同体に対する抗体またはS i r 2 3 0 4のヒト相同体に結合するオリゴマーもしくはポリマー；Kuを不活性化する剤、例えば抗Ku抗体またはKu結合オリゴマーもしくはポリマーである。非相同端連結を阻害する剤のいずれも、単独で投与してもよいし、または非相同端連結を阻害する1以上の他の剤と組み合わせて投与してもよい。

#### 【0008】

好ましい態様において、DNA配列は直鎖DNA配列である。好ましい態様に



において、直鎖DNA配列は、1以上の一本鎖オーバーハング（類）を有することが可能である。

#### 【0009】

好ましい態様において、選択されるDNA配列には、ターゲッティング配列が隣接する。ターゲッティング配列は、標的に相同であり、例えば標的DNAを改変しようとする部位または選択されるDNA配列を組み込もうとする部位に隣接するDNAに相同である。こうした隣接配列は、選択されるDNA配列の1以上の端、好ましくは両端に、存在することが可能である。2つの隣接配列が存在する場合、一方は標的の第一の領域に相同であるべきであり、そして他方は標的の第二の領域に相同であるべきである。

#### 【0010】

好ましい態様において、DNA配列は、1以上の突出一本鎖端を有し、例えば突出端の1つまたは両方が、3'端または5'端である。

好ましい態様において、相同組換えを増進する剤は：Rad52タンパク質または機能するその断片；Rad51タンパク質または機能するその断片；Rad54タンパク質または機能するその断片；あるいはそれらの組み合わせである。

#### 【0011】

好ましい態様において、相同組換えを増進する剤は、DNA配列に付着し、例えばDNA配列上にコーティングされる。好ましい態様において、Rad52タンパク質または機能するその断片は、DNA配列に付着し、例えばDNA配列上にコーティングされる。

#### 【0012】

好ましい態様において、Rad52タンパク質またはその断片はヒトRad52（hRad52）である。

好ましい態様において、抗Ku抗体は：抗Ku70抗体；抗Ku80抗体である。好ましい態様において、抗Ku抗体は：ヒト化抗体；ヒト抗体；抗体断片、例えばFab、Fab'、F(ab')<sup>2</sup>またはF(v)断片である。

#### 【0013】

好ましい態様において、少なくとも1つの抗Ku抗体は：DNA配列；Rad



52タンパク質またはその断片に共有結合している。別の好ましい態様において、少なくとも1つの抗Ku抗体は：DNA配列；Rad52タンパク質またはその断片に非共有結合している。

#### 【0014】

好ましい態様において、抗Ku70抗体および抗Ku80抗体は、例えば複合体の構成要素として提供される。

好ましい態様において、細胞は：真核細胞である。好ましい態様において、細胞は真菌、植物または動物起源、例えば脊椎動物起源のものである。好ましい態様において、細胞は：哺乳動物細胞、例えば初代または二次哺乳動物細胞、例えば線維芽細胞、造血幹細胞、筋芽細胞、ケラチン形成細胞、上皮細胞、内皮細胞、グリア細胞、神経細胞、血液の形成された要素を含む細胞、筋肉細胞およびこれらの体細胞の前駆体；形質転換または不死化細胞株である。好ましくは、細胞はヒト細胞である。本方法に有用な不死化ヒト細胞株の例には、限定されるわけではないが：Bowes黒色腫細胞(ATCC寄託番号CRL 9607)、Dauid細胞(ATCC寄託番号CCL 213)、HeLa細胞およびHeLa細胞派生物(derivative)(ATCC寄託番号CCL2 CCL2.1、およびCCL 2.2)、HL-60細胞(ATCC寄託番号CCL 240)、HT1080細胞(ATCC寄託番号CCL 121)、Jurkat細胞(ATCC寄託番号TIB 152)、KB癌腫細胞(ATCC寄託番号CCL 17)、K-562白血病細胞(ATCC寄託番号CCL 243)、MCF-7乳癌細胞(ATCC寄託番号BTH 22)、MOLT-4細胞(ATCC寄託番号1582)、Namalwa細胞(ATCC寄託番号CRL 1432)、Rafji細胞(ATCC寄託番号CCL 86)、RPMI 8226細胞(ATCC寄託番号CCL 155)、U-937細胞(ATCC寄託番号1593)、WI-28VA13下位株2R4細胞(ATCC寄託番号CCL 155)、CCRF-CEM細胞(ATCC寄託番号CCL 119)および2780AD卵巢癌細胞(Van Der Blickら, Cancer Res. 48:5927-5932, 1988)と共に、ヒト細胞および別の種の細胞の融合によって産生されるヘテロハイブリドーマ細胞が含まれる。別の



態様において、不死化細胞株は、ヒト細胞株以外の細胞株、例えばCHO細胞株、COS細胞株であってもよい。

#### 【0015】

好ましい態様において、構成要素、例えば複合体の構成要素は、マイクロインジェクションによって、細胞に導入される。

1つの好ましい態様において、選択されるDNA配列は、10、8、6、5、4、3、2未満、または1ヌクレオチド、例えば置換、または欠失、または挿入により、標的DNAと異なる。

#### 【0016】

好ましい態様において、標的DNAは突然変異を含み、例えば標的配列は、約10、8、6、5、4、3、2、または1ヌクレオチド、野生型配列と異なる。好ましくは、突然変異は点突然変異、例えば挿入、欠失または置換による突然変異である。

#### 【0017】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は、疾患または機能障害に関連し、例えば該疾患または機能障害を引き起こすか、それに貢献するか、それを左右するかまたは調節する。好ましくは、疾患または機能障害は：嚢胞性線維症；鎌形赤血球貧血；血友病A；血友病B；フォンウィルブラント疾患3型；色素性乾皮症；地中海貧血症；レッシューナイラン（Lesch-Nyhan）症候群；プロテインC耐性；リソソーム貯蔵疾患、例えばゴシェ病、ファブリー病；ムコ多糖症（MPS）1型（ハーレーシャイエ（Hurler-Scheie）症候群）、MPS II型（ハンター症候群）、MPS IIIA型（サンフィリオ（Sanfilippo）A症候群）、MPS IIIB型（サンフィリオB症候群）、MPS IIIC型（サンフィリオC症候群）、MPS IIID型（サンフィリオD症候群）、MPS IVA型（モルキオA症候群）、MPS IVB型（モルキオB症候群）、MPS VI型（マロトラーリー（Maroteaux-Lary）症候群）、MPS VII型（スライ症候群）である。

#### 【0018】



好ましい態様において、標的DNAは突然変異を含み、そして選択されるDNA配列は突然変異を訂正することが可能な正常野生型配列を含む。

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は嚢胞性線維症膜貫通制御因子（CFTR）遺伝子中にある。好ましくは、突然変異は、CFTRタンパク質コード領域のコドン508のアミノ酸を改変するものであり、例えば突然変異は、CFTRタンパク質のコドン508のフェニルアラニンを除去する、3塩基対インフレーム欠失である。CFTRタンパク質におけるフェニルアラニン508のこの欠失は、嚢胞性線維症を有する被験者に高い割合で見られる。したがって、好ましい態様において、野生型CFTR遺伝子に見られるようなフェニルアラニン508をコードする配列を含む、選択されるDNA配列を用いて、突然変異CFTR遺伝子を標的とし、そして訂正することが可能である。

#### 【0019】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異はヒトβ-グロビン遺伝子中にある。好ましくは、突然変異は、β-グロビン遺伝子の六番目のコドンのアミノ酸を改変するものであり、例えば突然変異は、β-グロビン遺伝子の六番目のコドンにおけるAからTの置換である。この突然変異は、β-グロビントタンパク質において、鎌形赤血球貧血を有する被験者に見られる、グルタミン酸からバリンへの変化を導く。したがって、好ましい態様において、コドン6で野生型アミノ酸残基をコードする、選択されるDNA、例えば野生型β-グロビン遺伝子の六番目のコドン内で見られるように、Aを含む、選択されるDNA配列を用いて、突然変異β-グロビン遺伝子を標的とし、そして訂正することが可能である。

#### 【0020】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は因子VII遺伝子中にある。例えば、突然変異は、因子VII遺伝子のエクソン23、24、および／またはエクソン25にある可能性がある。好ましくは、突然変異は、因子VIIタンパク質コード領域のコード領域のコドン2209でアミノ酸を改変するものであり、例えば突然変異は、因子VIIのアミノ酸2



209でアルギニンからグルタミンへの変化を導く、因子V I I I 遺伝子のエクソン24におけるGからAの置換である。好ましくは、突然変異は、因子V I I I タンパク質コード領域のコード領域のコドン2229でアミノ酸を改変するものであり、例えば突然変異は、因子V I I I のアミノ酸2229でトリプトファンからシステインへの変化を導く、因子V I I I 遺伝子のエクソン25におけるGからTの置換である。これらの突然変異は、中程度から重度の血友病Aと関連付けられてきている。したがって、好ましい態様において、因子V I I I 遺伝子のコード領域のコドン2209で野生型アミノ酸をコードするDNA、または因子V I I I 遺伝子のコード領域のコドン2229で野生型アミノ酸をコードするDNAいずれか、あるいは両方を含む、選択されるDNAの配列を用いて、突然変異因子V I I I 遺伝子を標的とし、そして訂正することが可能である。

#### 【0021】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は因子I X 遺伝子中にある。例えば、血友病Bを有する被験者において、突然変異の大部分は、因子I X 遺伝子における点突然変異である。したがって、好ましい態様において、選択されるDNA配列は、血友病Bに関連する因子I X 遺伝子中の1以上の点突然変異を標的とし、そして訂正するため、野生型因子I X 遺伝子由来の少なくとも1つのヌクレオチドを有する1以上のヌクレオチドを含むことが可能である。

#### 【0022】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異はフォンウィルブランド因子遺伝子中にある。好ましくは、突然変異は、フォンウィルブランド遺伝子のエクソン18における位2679-2684の6シトシン伸長中の単一シトシン欠失である。この突然変異は、フォンウィルブランド疾患3型を有する被験者に、かなりの割合で見られる。フォンウィルブランド疾患3型に関連する他の突然変異、例えば点突然変異もまた、本明細書に記載されるように、改変可能である。したがって、好ましい態様において、野生型フォンウィルブランド遺伝子に見られる配列、例えばフォンウィルブランド遺伝子の位2679-2684の6つのシトシンを含む、選択されるDNA配列を用いて、突然変異



フォンウィルブランド遺伝子を標的とし、そして訂正することが可能である。

#### 【0023】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は色素性乾皮症グループG (XP-G) 遺伝子中にある。好ましくは、突然変異は、XP-G 遺伝子に見られる245塩基対エクソンの位19-21の3アデニン伸長中の単一アデニン欠失である。この欠失は、色素性乾皮症を導く。したがって、好ましい態様において、XP-G 遺伝子の野生型配列、例えば245塩基対エクソンの位19-21の3つのアデニンを含む、選択されるDNAを用いて、突然変異XP-G 遺伝子を標的とし、そして訂正することが可能である。

#### 【0024】

好ましくは、Msh2、Msh6、Msh3、Mlh1、Pms2、Mlh3、Pms1などのミスマッチ修復タンパク質を不活性化する剤もまた、提供される。剤は、複合体中に含まれていてもよい。

#### 【0025】

別の好ましい態様において、改変は、選択されるDNA配列および標的DNA、例えば染色体間の相同組換えを含む。

好ましい態様において、選択されるDNA配列は、1より多いヌクレオチドで標的DNAと異なり、例えば標的、または選択されるDNA配列が、非対領域、例えばループアウト領域を有するように、十分な数のヌクレオチドで、標的と異なる。こうした適用において、Msh2、Msh6、Msh3、Mlh1、Pms2、Mlh3、Pms1もまた、例えば複合体の一部として提供してもよい。

#### 【0026】

好ましい態様において、改変は、標的DNAへの選択される配列の組込みを含み、そして選択されるDNAは、標的上のあらかじめ選択される要素と、あらかじめ選択される関係にあるように組み込まれ、例えば一方が制御要素であり、そして他方がタンパク質をコードする配列である場合、制御要素が、タンパク質コード配列の発現を制御するよう機能する。選択される組込みを促進する隣接配列を用いてもよい。選択されるDNA配列は、選択される標的配列、例えば遺伝子またはコード配列の5'、3'、または該配列内に組み込むことが可能である。



## 【0027】

好ましい態様において、改変は、選択されるDNA配列の組込みを含み、そして選択されるDNA配列は制御配列、例えば外因性制御配列である。好ましい態様において、制御配列は、1以上の：プロモーター、エンハンサー、上流活性化配列（UAS）、足場付着領域または転写因子結合部位を含む。好ましい態様において、制御配列は：メタロチオネインーI遺伝子、例えばマウスメタロチオネインーI遺伝子由来の制御配列、SV-40遺伝子由来の制御配列、サイトメガロウイルス遺伝子由来の制御配列、コラーゲン遺伝子由来の制御配列、アクチン遺伝子由来の制御配列、免疫グロブリン遺伝子由来の制御配列、HMG-CoAレダクターゼ遺伝子由来の制御配列、 $\gamma$ アクチン遺伝子由来の制御配列、転写活性化因子YY1遺伝子由来の制御配列、フィブロネクチン遺伝子由来の制御配列、またはEF-1 $\alpha$ 遺伝子由来の制御配列を含む。

## 【0028】

好ましい態様において、選択されるDNA配列はエクソンを含む。好ましくは、外因性エクソンは：CAP部位、ヌクレオチド配列ATG、および／または標的とされる内因性遺伝子とインフレームのコードDNAを含む。

## 【0029】

好ましい態様において、選択されるDNA配列はスプライスドナー部位を含む。

好ましい態様において、選択されるDNA配列は、標的に組み込まれた際、内因性コード配列を制御するよう作用する、外因性制御配列を含む。選択されるDNA配列は、標的における内因性遺伝子のコード領域の上流、または標的における内因性遺伝子の内因性制御配列の上流に組み込むことが可能である。別の好ましい態様において、選択されるDNA配列は、内因性遺伝子またはコード領域の下流、あるいはイントロンまたは内因性遺伝子内に組み込むことが可能である。別の好ましい態様において、選択されるDNA配列は、内因性遺伝子の内因性制御配列が不活性である、例えば完全にまたは部分的に欠失されるように、組み込むことが可能である。

## 【0030】



好ましい態様において、選択されるDNA配列は内因性遺伝子上流であり、そして内因性遺伝子の第二のエクソンに連結される。

好ましい態様において、内因性遺伝子は：ホルモン、サイトカイン、抗原、抗体、酵素、凝固因子、輸送タンパク質、受容体、制御タンパク質、構造タンパク質または転写因子をコードする。好ましい態様において、内因性遺伝子は、以下のタンパク質：エリスロポエチン、カルシトニン、成長ホルモン、インスリン、インスリノトロピン、インスリン様増殖因子、副甲状腺ホルモン、 $\alpha$ 2-インターフェロン (IFNA2)、 $\beta$ -インターフェロン、 $\gamma$ -インターフェロン、神経増殖因子類、FSH $\beta$ 、TGF- $\beta$ 、腫瘍壊死因子、グルカゴン、骨増殖因子-2、骨増殖因子-7、TSH- $\beta$ 、インターロイキン1、インターロイキン2、インターロイキン3、インターロイキン6、インターロイキン11、インターロイキン12、CSF-顆粒球 (GCSF)、CSF-マクロファージ、CSF-顆粒球/マクロファージ、免疫グロブリン類、触媒性抗体類、プロテインキナーゼC、グルコセレブロシダーゼ、スーパーオキシドジスムターゼ、組織プラスミノゲン活性化因子、ウロキナーゼ、アンチトロンビンIII、DNアーゼ、 $\alpha$ -ガラクトシダーゼ、チロシンヒドロキシラーゼ、血液凝固因子V、血液凝固因子VII、血液凝固因子VIII、血液凝固因子IX、血液凝固因子X、血液凝固因子XIII、アポリポタンパク質E、アポリポタンパク質A-I、グロビン類、低密度リポタンパク質受容体、IL-2受容体、IL-2アンタゴニスト類、 $\alpha$ -1-アンチトリプシン、免疫反応修飾剤類、 $\beta$ -グルコセラミダーゼ、 $\alpha$ -イズロニダーゼ、 $\alpha$ -L-イズロニダーゼ、グルコサミン-N-スルファターゼ、 $\alpha$ -N-アセチルグルコサミニダーゼ、アセチル補酵素A： $\alpha$ -グルコサミン-N-アセチルトランスフェラーゼ、N-アセチルグルコサミン-6-スルファターゼ、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、N-アセチルガラクトサミン-6-スルファターゼ、および可溶性CD4のいずれかをコードする。

### 【0031】

好ましい態様において、内因性遺伝子は卵胞刺激ホルモン $\beta$  (FSH $\beta$ ) をコードし、そして選択されるDNA配列は、制御配列、例えばFSH $\beta$  遺伝子の制



御配列と配列が異なる制御配列を含む。好ましくは、選択されるDNA配列には、ターゲッティング配列が隣接し、例えばこうしたターゲッティング配列は、選択されるDNA配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲッティング配列は、F S H  $\beta$  コード領域（配列番号1）の5'の領域に相同である。好ましい態様において、ターゲッティング配列は、F S H  $\beta$  コード配列内で、またはF S H  $\beta$  コード配列上流で、相同組換えを指示する。好ましい態様において、ターゲッティング配列は、ヒトF S H  $\beta$  配列のヌクレオチドー7454からー1417（番号付けは翻訳開始部位に対する）に対応する配列番号2由来、またはヒトF S H  $\beta$  配列のヌクレオチドー696からー155に対応する配列番号3由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。

#### 【0032】

好ましい態様において、内因性遺伝子はインターフェロン $\alpha$ 2（IFN $\alpha$ 2）をコードし、そして選択されるDNA配列は、制御配列、例えばIFN $\alpha$ 2遺伝子の制御配列と配列が異なる制御配列を含む。好ましくは、選択されるDNA配列には、ターゲッティング配列が隣接し、例えばこうしたターゲッティング配列は、選択されるDNA配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲッティング配列は、IFN $\alpha$ 2コード領域の5'の領域に相同である。好ましい態様において、ターゲッティング配列は、IFN $\alpha$ 2コード配列上流の領域内で、相同組換えを指示する。好ましい態様において、ターゲッティング配列は、ヒトIFN $\alpha$ 2配列のヌクレオチドー4074からー511（番号付けは翻訳開始部位に対する）に対応する配列番号4由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。例えば、該配列は：ヒトIFN $\alpha$ 2配列のヌクレオチドー4074からー3796に対応する配列番号7由来の少なくとも20、30、50、または100ヌクレオチド；ヒトIFN $\alpha$ 2配列のヌクレオチドー582からー510に対応する配列番号8由来の少なくとも20、30、または50ヌクレオチド；ヒトIFN $\alpha$ 2配列のヌクレオチドー3795からー583に対応する配列番号9由来の少なくとも20、30、50、100、または1000ヌクレオチドを含んでもよい。



## 【0033】

好ましい態様において、内因性遺伝子は顆粒球コロニー刺激因子（GCSF）をコードし、そして選択されるDNA配列は、制御配列、例えばGCSF遺伝子の制御配列と配列が異なる制御配列を含む。好ましくは、選択されるDNA配列には、ターゲティング配列が隣接し、例えばこうしたターゲティング配列は、選択されるDNA配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲティング配列は、GCSFコード領域の5'の領域に相同である。好ましい態様において、ターゲティング配列は、GCSFコード配列内で、またはGCSFコード配列上流で、相同組換えを指示する。好ましい態様において、ターゲティング配列は、ヒトGCSF配列のヌクレオチドー6, 578から101（番号付けは翻訳開始部位に対する）に対応する配列番号5由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。例えば、標的配列は、ヒトGCSF遺伝子のヌクレオチドー6, 578から-364に対応する配列番号6由来の20、30、50、100または1000ヌクレオチドを含んでもよい。

## 【0034】

別の好ましい態様において、DNA配列はコード領域を含み、例えば選択されるDNA配列はタンパク質をコードする。好ましい態様において、コード領域は：ホルモン、サイトカイン、抗原、抗体、酵素、凝固因子、輸送タンパク質、受容体、制御タンパク質、構造タンパク質または転写因子をコードする。好ましい態様において、コード領域は、以下のタンパク質：エリスロポエチン、カルシトニン、成長ホルモン、インスリン、インスリノトロピン、インスリン様増殖因子、副甲状腺ホルモン、 $\alpha$ 2-インターフェロン（IFNA2）、 $\beta$ -インターフェロン、 $\gamma$ -インターフェロン、神経増殖因子類、FSH $\beta$ 、TGF- $\beta$ 、腫瘍壊死因子、グルカゴン、骨増殖因子-2、骨増殖因子-7、TSH- $\beta$ 、インターロイキン1、インターロイキン2、インターロイキン3、インターロイキン6、インターロイキン11、インターロイキン12、CSF-顆粒球（GCSF）、CSF-マクロファージ、CSF-顆粒球/マクロファージ、免疫グロブリン類、触媒性抗体類、プロテインキナーゼC、グルコセレブロシダーゼ、スーパー



オキシドジスムターゼ、組織プラスミノーゲン活性化因子、ウロキナーゼ、アンチトロンビンⅠⅠⅠ、DNアーゼ、 $\alpha$ -ガラクトシダーゼ、チロシンヒドロキシラーゼ、血液凝固因子Ⅴ、血液凝固因子ⅤⅠⅠ、血液凝固因子ⅤⅠⅠⅠ、血液凝固因子ⅠⅩ、血液凝固因子Ⅹ、血液凝固因子ⅩⅠⅠⅠ、アポリポタンパク質Ⅴ、アポリポタンパク質Ⅰ、グロビン類、低密度リポタンパク質受容体、ⅠⅠ-Ⅱ受容体、ⅠⅠ-Ⅱアンタゴニスト類、 $\alpha$ -Ⅰ-アンチトリプシン、免疫反応修飾剤類、 $\beta$ -グルコセラミダーゼ、 $\alpha$ -イズロニダーゼ、 $\alpha$ -Ⅰ-イズロニダーゼ、グルコサミン-Ⅰ-スルファターゼ、 $\alpha$ -Ⅰ-アセチルグルコサミニダーゼ、アセチル補酵素Ⅰ： $\alpha$ -グルコサミン-Ⅰ-アセチルトランスフェラーゼ、Ⅰ-アセチルグルコサミン-Ⅰ-スルファターゼ、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、Ⅰ-アセチルガラクトサミン-Ⅰ-スルファターゼ、および可溶性CD4のいずれかをコードする。

#### 【0035】

好ましい態様において、選択されるDNA配列が内因性制御要素の調節下にあるように、該配列を標的内に組み込むことが可能である。選択されるDNAは、内因性制御配列の下流または内因性遺伝子のコード領域の上流および遺伝子の内因性制御配列の下流に組み込むことが可能である。別の好ましい態様において、選択されるDNAは、内因性遺伝子のコード領域が不活性化される、例えば完全にまたは部分的に欠失されるように、内因性制御配列の下流に組み込むことが可能である。

#### 【0036】

好ましい態様において、該方法は、ミスマッチ修復タンパク質、例えばMsh2、Msh6、Msh3、Mlh1、Pms2、Mlh3、Pms1、または他のミスマッチ修復タンパク質、あるいはそれらの組み合わせを阻害する剤を導入することをさらに含む。好ましくは、剤は、ミスマッチ修復タンパク質の発現を阻害する剤であり、例えば剤はアンチセンスRNAである。好ましい態様において、剤はミスマッチ修復タンパク質に対する抗体である。好ましい態様において、ミスマッチ修復タンパク質に対する抗体は、複合体に共有または非共有結合している。



## 【0037】

別の側面において、本発明は、選択されるDNA配列、例えば本明細書記載の選択されるDNA配列をテンプレートとして用いて、標的DNA、例えば染色体、例えば本明細書記載の標的DNAでの改変を促進するための組成物、例えば構成要素の複合体を特徴とする。組成物は：（a）選択されるDNA配列を含む二本鎖DNA配列；（b）相同組換えを増進する剤、例えばRad52タンパク質または機能するその断片；および（c）非相同端連結を阻害する剤、例えばKuを不活性化する剤を含む。該組成物を用いて、例えば組込みによって標的DNA配列を改変可能である。

## 【0038】

好ましい態様において、非相同端連結を阻害する剤は：hMre11を不活性化する剤、例えば抗hMre11抗体またはhMre11結合オリゴマーもしくはポリマー；hRad50を不活性化する剤、例えば抗hRad50抗体またはhRad50結合オリゴマーもしくはポリマー；Nbs1を不活性化する剤、例えば抗Nbs1抗体またはhNbs1結合オリゴマーもしくはポリマー；ヒトリガーゼ4（hLig4）を不活性化する剤、例えば抗hLig4抗体またはhLig4結合オリゴマーもしくはポリマー；hXrcc4を不活性化する剤、例えば抗hXrcc4抗体またはhXrcc4結合オリゴマーもしくはポリマー；Rap1のヒト相同体を不活性化する剤、例えばRap1のヒト相同体に対する抗体またはRap1のヒト相同体に結合するオリゴマーもしくはポリマー；Sir2304のヒト相同体を不活性化する剤、例えばSir2304のヒト相同体に対する抗体またはSir2304のヒト相同体に結合するオリゴマーもしくはポリマー；Kuを不活性化する剤、例えば抗Ku抗体またはKu結合オリゴマーもしくはポリマーである。非相同端連結を阻害する剤のいずれも、単独で投与してもよいし、または非相同端連結を阻害する1以上の他の剤と組み合わせて投与してもよい。

## 【0039】

好ましい態様において、DNA配列は直鎖DNA配列である。好ましい態様において、直鎖DNA配列は、1以上の一本鎖オーバーハング（類）を有すること



が可能である。

#### 【0040】

好ましい態様において、選択されるDNA配列にはターゲティング配列が隣接する。ターゲティング配列は、標的に相同であり、例えば標的DNAを改変しようとする部位または選択されるDNA配列を挿入しようとする部位に隣接するDNAに相同である。こうした隣接配列は、選択されるDNA配列の1以上の端、好ましくは両端に、存在することが可能である。2つの隣接配列が存在する場合、一方は標的の第一の領域に相同であるべきであり、そして他方は標的の第二の領域に相同であるべきである。

#### 【0041】

好ましい態様において、DNA配列は、1以上の突出一本鎖端を有し、例えば突出端の1つまたは両方が、3'端または5'端である。

好ましい態様において、相同組換えを増進する剤は：Rad52タンパク質または機能するその断片；Rad51タンパク質または機能するその断片；Rad54タンパク質または機能するその断片；あるいはそれらの組み合わせである。

#### 【0042】

好ましい態様において、相同組換えを増進する剤は、DNA配列に付着し、例えばDNA配列上にコーティングされる。好ましい態様において、Rad52タンパク質または機能するその断片は、選択されるDNA配列に付着し、例えば選択されるDNA配列上にコーティングされる。

#### 【0043】

好ましい態様において、Rad52タンパク質またはその断片はヒトRad52 (hRad52) である。

好ましい態様において、抗Ku抗体は：抗Ku70抗体；抗Ku80抗体である。好ましい態様において、抗Ku抗体は：ヒト化抗体；ヒト抗体；抗体断片、例えばFab、Fab'、F(ab')<sup>2</sup> またはF(v)断片である。

#### 【0044】

好ましい態様において、少なくとも1つの抗Ku抗体は：選択されるDNA配列；Rad52タンパク質またはその断片に共有結合している。別の好ましい態



様において、少なくとも1つの抗Ku抗体は：選択されるDNA配列；Rad52タンパク質またはその断片に非共有結合している。

#### 【0045】

好ましい態様において、組成物は、抗Ku70抗体および抗Ku80抗体を含む。

好ましい態様において、選択されるDNA配列は、10、8、6、5、4、3、2未満、または1ヌクレオチド、例えば置換、または欠失、または挿入により、標的DNAと異なる。

#### 【0046】

好ましい態様において、標的DNAは突然変異を含み、例えば標的配列は、約10、8、6、5、4、3、2、または1ヌクレオチド、野生型配列と異なる。好ましくは、突然変異は点突然変異、例えば挿入、欠失または置換による突然変異である。

#### 【0047】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は、疾患または機能障害に関連し、例えば該疾患または機能障害を引き起こすか、それに貢献するか、それを左右するかまたは調節する。好ましくは、疾患または機能障害は：嚢胞性線維症；鎌形赤血球貧血；血友病A；血友病B；フォンウィルブランド疾患3型；色素性乾皮症；地中海貧血症；レッシューナイラン症候群；プロテインC耐性；リソソーム貯蔵疾患、例えばゴシェ病、ファブリー病；ムコ多糖症（MPS）1型（ハーレーーシャイエ症候群）、MPS II型（ハンター症候群）、MPS IIIA型（サンフィリオA症候群）、MPS IIIB型（サンフィリオB症候群）、MPS IIIC型（サンフィリオC症候群）、MPS IIID型（サンフィリオD症候群）、MPS IVA型（モルキオA症候群）、MPS IVB型（モルキオB症候群）、MPS VI型（マロトラーラー症候群）、MPS VII型（スライ症候群）である。

#### 【0048】

好ましい態様において、標的DNAは突然変異を含み、そして選択されるDNA配列は突然変異を訂正することが可能な正常野生型配列を含む。



好ましい態様において、標的DNAは突然変異を含み、そして突然変異は嚢胞性線維症膜貫通制御因子（CFTR）遺伝子中にある。好ましくは、突然変異は、CFTRタンパク質コード領域のコドン508のアミノ酸を改変するものであり、例えば突然変異は、CFTRタンパク質のコドン508のフェニルアラニンを除去する、3塩基対インフレーム欠失である。CFTRタンパク質におけるフェニルアラニン508のこの欠失は、嚢胞性線維症を有する被験者に高い割合で見られる。したがって、好ましい態様において、野生型CFTR遺伝子に見られるようなフェニルアラニン508をコードする配列を含む、選択されるDNA配列を用いて、突然変異CFTR遺伝子を標的とし、そして訂正することが可能である。

#### 【0049】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異はヒトβ-グロビン遺伝子中にある。好ましくは、突然変異は、β-グロビン遺伝子の六番目のコドンのアミノ酸を改変するものであり、例えば突然変異は、β-グロビン遺伝子の六番目のコドンにおけるAからTの置換である。この突然変異は、β-グロビントタンパク質において、鎌形赤血球貧血を有する被験者に見られる、グルタミン酸からバリンへの変化を導く。したがって、好ましい態様において、コドン6で野生型アミノ酸残基をコードする、選択されるDNA、例えば野生型β-グロビン遺伝子の六番目のコドン内で見られるように、Aを含む、選択されるDNA配列を用いて、突然変異β-グロビン遺伝子を標的とし、そして訂正することが可能である。

#### 【0050】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は因子VII遺伝子中にある。例えば、突然変異は、因子VII遺伝子のエクソン23、24、および／またはエクソン25にある可能性がある。好ましくは、突然変異は、因子VIIタンパク質コード領域のコード領域のコドン2209でアミノ酸を改変するものであり、例えば突然変異は、因子VIIのアミノ酸2209でアルギニンからグルタミンへの変化を導く、因子VII遺伝子のエクソン24におけるGからAの置換である。好ましくは、突然変異は、因子VII



I タンパク質コード領域のコード領域のコドン2229でアミノ酸を改変するものであり、例えば突然変異は、因子V I I Iのアミノ酸2229でトリプトファンからシステインへの変化を導く、因子V I I I遺伝子のエクソン25におけるGからTの置換である。これらの突然変異は、中程度から重度の血友病Aと関連付けられてきている。したがって、好ましい態様において、因子V I I I遺伝子のコード領域のコドン2209で野生型アミノ酸をコードするDNA、または因子V I I I遺伝子のコード領域のコドン2229で野生型アミノ酸をコードするDNAいずれか、あるいは両方を含む、選択されるDNA配列を用いて、突然変異因子V I I I遺伝子を標的とし、そして訂正することが可能である。

#### 【0051】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は因子I X遺伝子中にある。例えば、血友病Bを有する被験者において、突然変異の大部分は、因子I X遺伝子における点突然変異である。したがって、好ましい態様において、選択されるDNA配列は、血友病Bに関連する因子I X遺伝子中の1以上の点突然変異を標的とし、そして訂正するため、野生型因子I X遺伝子由来の少なくとも1つのヌクレオチドを有する1以上のヌクレオチドを含むことが可能である。

#### 【0052】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異はフォンウィルブランド因子遺伝子中にある。好ましくは、突然変異は、フォンウィルブランド遺伝子のエクソン18における位2679-2684の6シトシン伸長中の単一シトシン欠失である。この突然変異は、フォンウィルブランド疾患3型を有する被験者に、かなりの割合で見られる。フォンウィルブランド疾患3型に関連する他の突然変異、例えば点突然変異もまた、本明細書に記載されるように、改変可能である。したがって、好ましい態様において、野生型フォンウィルブランド遺伝子に見られる配列、例えばフォンウィルブランド遺伝子のエクソン18の位2679-2684の6つのシトシンを含む、選択されるDNA配列を用いて、突然変異フォンウィルブランド遺伝子を標的とし、そして訂正することが可能である。



## 【0053】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は色素性乾皮症グループG (XP-G) 遺伝子中にある。好ましくは、突然変異は、XP-G 遺伝子に見られる245塩基対エクソンの位19-21の3アデニン伸長中の単一アデニン欠失である。この欠失は、色素性乾皮症を導く。したがって、好ましい態様において、XP-G 遺伝子の野生型配列、例えばXP-G 遺伝子の245塩基対エクソンの位19-21の3つのアデニンを含む、選択されるDNAを用いて、突然変異XP-G 遺伝子を標的とし、そして訂正することが可能である。

## 【0054】

別の好ましい態様において、選択されるDNA配列は、1より多いヌクレオチドで標的DNAと異なり、例えば標的、または選択されるDNA配列が、非対領域、例えばループアウト領域を有するように、十分な数のヌクレオチドで、標的と異なる。好ましくは、Msh2、Msh6、Msh3、Mlh1、Pms2、Mlh3、Pms1、またはそれらの組み合わせなどのミスマッチ修復タンパク質を不活性化する剤もまた、組成物に含んでもよく、例えば該剤を複合体に含んでもよい。

## 【0055】

好ましい態様において、選択されるDNA配列は、標的DNA上のあらかじめ選択される要素と、あらかじめ選択される関係にあるように組み込まれることが可能であるように、隣接配列を有する。例えば、選択されるDNAが制御配列であり、そして標的DNAがタンパク質をコードする場合、隣接配列は、制御要素がタンパク質コード配列の発現を制御するよう機能するように、組み込むであろうものである。選択される組込みを促進する隣接配列を用いることが可能である。選択されるDNA配列は、選択される標的配列、例えば標的中の遺伝子またはコード領域の5'、3'、または該領域内に組み込まれることが可能であるような隣接配列を有することが可能である。

## 【0056】

好ましい態様において、選択されるDNA配列は、制御配列、例えば外因性制



御配列を含む。好ましい態様において、制御配列は、1以上の：プロモーター、エンハンサー、UAS、足場付着領域または転写因子結合部位を含む。好ましい態様において、制御配列は：メタロチオネインーI遺伝子、例えばマウスメタロチオネインーI遺伝子由来の制御配列、SV-40遺伝子由来の制御配列、サイトメガロウイルス遺伝子由来の制御配列、コラーゲン遺伝子由来の制御配列、アクチン遺伝子由来の制御配列、免疫グロブリン遺伝子由来の制御配列、HMG-CoAレダクターゼ遺伝子由来の制御配列、 $\gamma$ アクチン遺伝子由来の制御配列、転写活性化因子YY1遺伝子由来の制御配列、フィブロネクチン遺伝子由来の制御配列、またはEF-1 $\alpha$ 遺伝子由来の制御配列を含む。

#### 【0057】

好ましい態様において、選択されるDNA配列はエクソンを含む。好ましくは、外因性エクソンは：CAP部位、ヌクレオチド配列ATG、および／または標的とされる内因性遺伝子とインフレームのコードDNAを含む。

#### 【0058】

好ましい態様において、選択されるDNA配列はスプライスドナー部位を含む。

好ましい態様において、外因性制御配列を有する、選択されるDNA配列を含む組成物は、選択されるDNA配列が、内因性配列の発現を制御するよう機能するように、標的内に組み込まれるような、隣接配列を有してもよい。選択されるDNAは、標的において、内因性遺伝子またはコード配列のコード領域上流の標的に組み込むことが可能であるし、あるいは標的において、内因性遺伝子またはコード配列の内因性制御配列上流の標的に組み込むことが可能である。別の好ましい態様において、選択されるDNA配列は、内因性遺伝子の内因性制御配列が不活性である、例えば完全にまたは部分的に欠失されるように、標的中に組み込むことが可能である。選択されるDNA配列は、内因性遺伝子またはコード領域の下流の標的に組み込むことが可能であるし、あるいは内因性遺伝子のイントロン内に組み込むことが可能である。

#### 【0059】

好ましい態様において、選択されるDNA配列は、制御配列、例えばFISH $\beta$



遺伝子の制御配列と配列が異なる制御配列を含む。好ましくは、選択されるDNA配列には、ターゲッティング配列が隣接し、例えばこうしたターゲッティング配列は、選択されるDNA配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲッティング配列は、F S H  $\beta$  コード領域（配列番号1）の5'の領域に相同である。好ましい態様において、ターゲッティング配列は、F S H  $\beta$  コード配列内で、またはF S H  $\beta$  コード配列上流で、相同組換えを指示する。好ましい態様において、ターゲッティング配列は、ヒトF S H  $\beta$  配列のヌクレオチド-7454から-1417（番号付けは翻訳開始部位に対する）に対応する配列番号2由来、またはヒトF S H  $\beta$  配列のヌクレオチド-696から-155に対応する配列番号3由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。

#### 【0060】

好ましい態様において、選択されるDNA配列は、制御配列、例えばI F N  $\alpha$  2 遺伝子の制御配列と配列が異なる制御配列を含む。好ましくは、選択されるDNA配列には、ターゲッティング配列が隣接し、例えばこうしたターゲッティング配列は、選択されるDNA配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲッティング配列は、I F N  $\alpha$  2 コード領域の5'の領域に相同である。好ましい態様において、ターゲッティング配列は、I F N  $\alpha$  2 コード配列上流の領域内で、相同組換えを指示する。好ましい態様において、ターゲッティング配列は、ヒトI F N  $\alpha$  2 配列のヌクレオチド-4074から-511（番号付けは翻訳開始部位に対する）に対応する配列番号4由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。例えば、該配列は：ヒトI F N  $\alpha$  2 配列のヌクレオチド-4074から-3796に対応する配列番号7由来の少なくとも20、30、50、または100ヌクレオチド；ヒトI F N  $\alpha$  2 配列のヌクレオチド-582から-510に対応する配列番号8由来の少なくとも20、30、または50ヌクレオチド；ヒトI F N  $\alpha$  2 配列のヌクレオチド-3795から-583に対応する配列番号9由来の少なくとも20、30、50、100、または1000ヌクレオチドを含んでもよい。



## 【0061】

好ましい態様において、選択されるDNA配列は、制御配列、例えばGCSF遺伝子の制御配列と配列が異なる制御配列を含む。好ましくは、選択されるDNA配列には、ターゲティング配列が隣接し、例えばこうしたターゲティング配列は、選択されるDNA配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲティング配列は、GCSFコード領域の5'の領域に相同である。好ましい態様において、ターゲティング配列は：GCSFコード配列内で；またはGCSFコード配列上流で、相同組換えを指示する。好ましい態様において、ターゲティング配列は、ヒトGCSF配列のヌクレオチド6, 578から101（番号付けは翻訳開始部位に対する）に対応する配列番号5由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。例えば、標的配列は、ヒトGCSF遺伝子のヌクレオチド6, 578から-364（番号付けは翻訳開始部位に対する）に対応する配列番号6由来の20、30、50、100または1000ヌクレオチドを含んでもよい。

## 【0062】

別の好ましい態様において、DNA配列はコード領域を含み、例えばDNA配列はタンパク質をコードする。好ましい態様において、コード領域は：ホルモン、サイトカイン、抗原、抗体、酵素、凝固因子、輸送タンパク質、受容体、制御タンパク質、構造タンパク質または転写因子をコードする。好ましい態様において、コード領域は、以下のタンパク質：エリスロポエチン、カルシトニン、成長ホルモン、インスリン、インスリノトロピン、インスリン様増殖因子、副甲状腺ホルモン、 $\alpha 2$ -インターフェロン（IFNA2）、 $\beta$ -インターフェロン、 $\gamma$ -インターフェロン、神経増殖因子類、FSH $\beta$ 、TGF- $\beta$ 、腫瘍壊死因子、グルカゴン、骨増殖因子-2、骨増殖因子-7、TSH- $\beta$ 、インターロイキン1、インターロイキン2、インターロイキン3、インターロイキン6、インターロイキン11、インターロイキン12、CSF-顆粒球（GCSF）、CSF-マクロファージ、CSF-顆粒球/マクロファージ、免疫グロブリン類、触媒性抗体類、プロテインキナーゼC、グルコセレブロシダーゼ、スーパーオキシドジスムターゼ、組織プラスミノゲン活性化因子、ウロキナーゼ、アンチトロンビ



ン I I I、D N アーゼ、 $\alpha$ -ガラクトシダーゼ、チロシンヒドロキシラーゼ、血液凝固因子 V、血液凝固因子 V I I、血液凝固因子 V I I I、血液凝固因子 I X、血液凝固因子 X、血液凝固因子 X I I I、アポリポタンパク質 E、アポリポタンパク質 A-I、グロビン類、低密度リポタンパク質受容体、I L-2 受容体、I L-2 アンタゴニスト類、 $\alpha$ -1-アンチトリプシン、免疫反応修飾剤類、 $\beta$ -グルコセラミダーゼ、 $\alpha$ -イズロニダーゼ、 $\alpha$ -L-イズロニダーゼ、グルコサミン-N-スルファターゼ、 $\alpha$ -N-アセチルグルコサミニダーゼ、アセチル補酵素 A： $\alpha$ -グルコサミン-N-アセチルトランスフェラーゼ、N-アセチルグルコサミン-6-スルファターゼ、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、N-アセチルガラクトサミン-6-スルファターゼ、および可溶性 C D 4 のいずれかをコードする。

#### 【0063】

好ましい態様において、選択される D N A 配列が、標的内に組み込まれる際、内因性制御要素の調節下にあるように、隣接配列を有してもよい。選択される D N A は、内因性制御配列の下流または内因性遺伝子のコード領域の上流および遺伝子の内因性制御配列の下流に組み込むことが可能である。別の好ましい態様において、選択される D N A は、内因性遺伝子のコード領域が不活性化される、例えば完全にまたは部分的に欠失されるように、内因性制御配列の下流に組み込むことが可能である。

#### 【0064】

好ましい態様において、組成物、例えば複合体は、細胞に導入する。好ましくは、細胞は真核細胞である。好ましい態様において、細胞は真菌、植物または動物起源、例えば脊椎動物起源のものである。好ましい態様において、細胞は：哺乳動物細胞、例えば初代または二次哺乳動物細胞、例えば線維芽細胞、造血幹細胞、筋芽細胞、ケラチン形成細胞、上皮細胞、内皮細胞、グリア細胞、神経細胞、血液の形成された要素を含む細胞、筋肉細胞およびこれらの体細胞の前駆体；形質転換または不死化細胞株である。好ましくは、細胞はヒト細胞である。本方法に有用な不死化ヒト細胞株の例には、限定されるわけではないが：B o w e s 黒色腫細胞 (A T C C 寄託番号 C R L 9607)、D a u d i 細胞 (A T C C



寄託番号CCL 213)、HeLa細胞およびHeLa細胞派生物(ATCC寄託番号CCL2 CCL2. 1、およびCCL 2. 2)、HL-60細胞(ATCC寄託番号CCL 240)、HT1080細胞(ATCC寄託番号CCL 121)、Jurkat細胞(ATCC寄託番号TIB 152)、KB癌腫細胞(ATCC寄託番号CCL 17)、K-562白血病細胞(ATCC寄託番号CCL 243)、MCF-7乳癌細胞(ATCC寄託番号BTH 22)、MOLT-4細胞(ATCC寄託番号1582)、Namalwa細胞(ATCC寄託番号CRL 1432)、Rafji細胞(ATCC寄託番号CCL 86)、RPMI 8226細胞(ATCC寄託番号CCL 155)、U-937細胞(ATCC寄託番号1593)、WI-28VA13下位株2R4細胞(ATCC寄託番号CLL 155)、CCRF-CEM細胞(ATCC寄託番号CCL 119)および2780AD卵巢癌細胞(Van Der Bliekら, Cancer Res. 48:5927-5932, 1988)と共に、ヒト細胞および別の種の細胞の融合によって産生されるヘテロハイブリドーマ細胞が含まれる。別の態様において、不死化細胞株は、ヒト細胞株以外の細胞株、例えばCHO細胞株、COS細胞株であってもよい。

#### 【0065】

好ましい態様において、組成物は、ミスマッチ修復タンパク質、例えばMsh2、Msh6、Msh3、Mlh1、Pms2、Mlh3、Pms1、または他のミスマッチ修復タンパク質、あるいはそれらの組み合わせを阻害する剤をさらに含む。好ましくは、剤は、ミスマッチ修復タンパク質の発現を阻害する剤であり、例えば剤はアンチセンスRNAである。好ましい態様において、剤はミスマッチ修復タンパク質に対する抗体である。好ましい態様において、ミスマッチ修復タンパク質に対する抗体は、組成物の1つ以上の構成要素に共有または非共有結合している。

#### 【0066】

別の側面において、本発明は、タンパク質を提供する方法を特徴とする。該方法は：本明細書に記載される方法によって作成された細胞を提供し、そして細胞がタンパク質を発現することを可能にすることを含む。



## 【0067】

別の好ましい態様において：該方法は：改変のため、標的とされる部位に、以下の構成要素：（a）選択されるDNA配列を含む二本鎖DNA配列；（b）相同組換えを増進する剤、例えばRad52タンパク質または機能するその断片；および（c）非相同端連結を阻害する剤、例えばKu不活性化剤が導入されている細胞を提供し、そして細胞がタンパク質を発現することを可能にすることを含む。タンパク質の発現は、例えば、DNAにコードされるタンパク質の発現を可能にすることによって、またはタンパク質の発現を活性化することによって、起こることが可能である。

## 【0068】

好ましい態様において、部位の改変、例えば選択されるDNA配列および標的DNA間の相同組換えまたは遺伝子訂正が、供給される相同組換え増進剤および非相同端連結阻害剤の非存在下で起こるであろう率より高い率で起こるために、相同組換えを増進する剤の濃度および非相同端連結を阻害する剤の濃度が、選択されるDNA配列および標的DNA間の相互作用部位で、十分であるように、構成要素（a）、（b）、および（c）が提供され、例えば細胞に導入される。非相同端連結を阻害する剤は、好ましくは、局所に提供される。

## 【0069】

好ましい態様において、構成要素（a）、（b）、および（c）は、共に導入してもよいし、または別個に導入してもよい。さらに、構成要素の2つを共に導入し、そして第三の構成要素を別個に導入してもよい。例えば、DNA配列および相同組換えを増進する剤、例えばRad52を共に導入してもよいし、またはDNA配列および非相同端連結を阻害する剤、例えばKu不活性化剤を共に導入してもよい。別の好ましい態様において、相同組換えを増進する剤および非相同端連結を阻害する剤を共に導入してもよい。

## 【0070】

構成要素の2つ、または好ましくはすべてを、複合体として提供してもよい。好ましい態様において、該方法は、標的DNAと：（a）選択されるDNA配列を含む二本鎖DNA配列；（b）相同組換えを増進する剤、例えばRad52タ



ンパク質または機能するその断片；および（c）非相同端連結を阻害する剤、例えばKuを不活性化する剤を含む複合体とを、例えば該複合体を細胞内に導入することによって接触させることを含む。

#### 【0071】

好ましい態様において、1つ、またはそれ以上、好ましくはすべての構成要素は、局所搬送、例えばマイクロインジェクションによって提供され、そして標的ゲノムまたは他の核酸から発現されない。特に好ましい態様において、非相同端連結を阻害する剤、例えば抗Ku抗体などのKu不活性化剤は、局所搬送、例えばマイクロインジェクションによって提供され、そして標的ゲノムまたは他の核酸から発現されない。

#### 【0072】

好ましい態様において、非相同端連結を阻害する剤は：hMr e 1 1を不活性化する剤、例えば抗hMr e 1 1抗体またはhMr e 1 1結合オリゴマーもしくはポリマー；hR a d 5 0を不活性化する剤、例えば抗hR a d 5 0抗体またはhR a d 5 0結合オリゴマーもしくはポリマー；Nb s 1を不活性化する剤、例えば抗Nb s 1抗体またはhNb s 1結合オリゴマーもしくはポリマー；ヒトリガーゼ4（hL i g 4）を不活性化する剤、例えば抗hL i g 4抗体またはhL i g 4結合オリゴマーもしくはポリマー；hX r c c 4を不活性化する剤、例えば抗hX r c c 4抗体またはhX r c c 4結合オリゴマーもしくはポリマー；Rap 1のヒト相同体を不活性化する剤、例えばRap 1のヒト相同体に対する抗体またはRap 1のヒト相同体に結合するオリゴマーもしくはポリマー；Sir 2 3 0 4のヒト相同体を不活性化する剤、例えばSir 2 3 0 4のヒト相同体に対する抗体またはSir 2 3 0 4のヒト相同体に結合するオリゴマーもしくはポリマー；Kuを不活性化する剤、例えば抗Ku抗体またはKu結合オリゴマーもしくはポリマーである。非相同端連結を阻害する剤のいずれも、単独で投与してもよいし、または非相同端連結を阻害する1以上の他の剤と組み合わせて投与してもよい。

#### 【0073】

好ましい態様において、DNA配列は直鎖DNA配列である。好ましい態様に



において、直鎖DNA配列は、1以上の一本鎖オーバーハング（類）を有することが可能である。

#### 【0074】

好ましい態様において、選択されるDNA配列にターゲティング配列が隣接する。ターゲティング配列は、標的に相同であり、例えば標的DNAを改変しようとする部位または選択されるDNA配列を組み込もうとする部位に隣接するDNAに相同である。こうした隣接配列は、選択されるDNA配列の1以上の端、好ましくは両端に、存在することが可能である。2つの隣接配列が存在する場合、一方は標的の第一の領域に相同であるべきであり、そして他方は標的の第二の領域に相同であるべきである。

#### 【0075】

好ましい態様において、DNA配列は、1以上の突出一本鎖端を有し、例えば突出端の1つまたは両方が、3'端または5'端である。

好ましい態様において、相同組換えを増進する剤は：Rad52タンパク質または機能するその断片；Rad51タンパク質または機能するその断片；Rad54タンパク質または機能するその断片；あるいはそれらの組み合わせである。

#### 【0076】

好ましい態様において、相同組換えを増進する剤は、DNA配列に付着し、例えばDNA配列上にコーティングされる。好ましい態様において、Rad52タンパク質または機能するその断片は、選択されるDNA配列に付着し、例えば選択されるDNA配列上にコーティングされる。

#### 【0077】

好ましい態様において、Rad52タンパク質またはその断片はヒトRad52（hRad52）である。

好ましい態様において、抗Ku抗体は：抗Ku70抗体；抗Ku80抗体である。好ましい態様において、抗Ku抗体は：ヒト化抗体；ヒト抗体；抗体断片、例えばFab、Fab'、F(ab')<sup>2</sup>またはF(v)断片である。

#### 【0078】

好ましい態様において、少なくとも1つの抗Ku抗体は：選択されるDNA配



列；相同組換えを増進する剤、例えばRad52タンパク質またはその断片に共有結合している。別の好ましい態様において、少なくとも1つの抗Ku抗体は：選択されるDNA配列；相同組換えを増進する剤、例えばRad52タンパク質またはその断片に非共有結合している。

### 【0079】

好ましい態様において、複合体は、例えば複合体の構成要素として提供される、抗Ku70抗体および抗Ku80抗体を含む。

好ましい態様において、細胞は：真核細胞である。好ましい態様において、細胞は真菌、植物または動物起源、例えば脊椎動物起源のものである。好ましい態様において、細胞は：哺乳動物細胞、例えば初代または二次哺乳動物細胞、例えば線維芽細胞、造血幹細胞、筋芽細胞、ケラチン形成細胞、上皮細胞、内皮細胞、グリア細胞、神経細胞、血液の形成された要素を含む細胞、筋肉細胞およびこれらの体細胞の前駆体；形質転換または不死化細胞株である。好ましくは、細胞はヒト細胞である。本方法に有用な不死化ヒト細胞株の例には、限定されるわけではないが：Bowes黒色腫細胞(ATCC寄託番号CRL 9607)、Dauidi細胞(ATCC寄託番号CCL 213)、HeLa細胞およびHeLa細胞派生物(ATCC寄託番号CCL2 CCL2. 1、およびCCL 2. 2)、HL-60細胞(ATCC寄託番号CCL 240)、HT1080細胞(ATCC寄託番号CCL 121)、Jurkat細胞(ATCC寄託番号TIB 152)、KB癌腫細胞(ATCC寄託番号CCL 17)、K-562白血病細胞(ATCC寄託番号CCL 243)、MCF-7乳癌細胞(ATCC寄託番号BTH 22)、MOLT-4細胞(ATCC寄託番号1582)、Namalwa細胞(ATCC寄託番号CRL 1432)、Raji細胞(ATCC寄託番号CCL 86)、RPMI 8226細胞(ATCC寄託番号CCL 155)、U-937細胞(ATCC寄託番号1593)、WI-28VA13下位株2R4細胞(ATCC寄託番号CCL 155)、CCRF-CEM細胞(ATCC寄託番号CCL 119)および2780AD卵巢癌細胞(Van Der Blickら, Cancer Res. 48:5927-5932, 1988)と共に、ヒト細胞および別の種の細胞の融合によって産



生されるヘテロハイブリドーマ細胞が含まれる。別の態様において、不死化細胞株は、ヒト細胞株以外の細胞株、例えばCHO細胞株、COS細胞株であってもよい。

#### 【0080】

好ましい態様において、構成要素、例えば複合体の構成要素は、マイクロインジェクションによって、細胞に導入される。

好ましい態様において、方法は、ミスマッチ修復タンパク質、例えばMsh2、Msh6、Msh3、Mlh1、Pms2、Mlh3、Pms1、または他のミスマッチ修復タンパク質、あるいはそれらの組み合わせを阻害する剤を導入することをさらに含む。好ましくは、剤は、ミスマッチ修復タンパク質の発現を阻害する剤であり、例えば剤はアンチセンスRNAである。好ましい態様において、剤はミスマッチ修復タンパク質に対する抗体である。好ましい態様において、ミスマッチ修復タンパク質に対する抗体は、複合体に共有または非共有結合している。

#### 【0081】

好ましい態様において、タンパク質は*in vitro*で発現される。他の好ましい態様において、細胞は被験者、例えばヒトにおいて提供され、そしてタンパク質は被験者において発現される。好ましい態様において、タンパク質は被験者において発現され、そして細胞は自己、同種異系、異種である。選択されるDNAは、*in vivo*で細胞に導入してもよいし、または細胞を被験者から除去し、選択されるDNAを*ex vivo*で導入し、そして細胞を被験者に戻してもよい。

#### 【0082】

好ましい態様において、選択されるDNA配列は、10、8、6、5、4、3、2未満、または1ヌクレオチド、例えば置換、または欠失、または挿入により、標的DNAと異なる。

#### 【0083】

好ましい態様において、標的DNAは突然変異を含み、例えば標的配列は、約10、8、6、5、4、3、2、または1ヌクレオチド、野生型配列と異なる。



好ましくは、突然変異は点突然変異、例えば挿入、欠失または置換による突然変異である。

#### 【0084】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は、疾患または機能障害に関連し、例えば該疾患または機能障害を引き起こすか、それに貢献するか、それを左右するかまたは調節する。好ましくは、疾患または機能障害は：嚢胞性線維症；鎌形赤血球貧血；血友病A；血友病B；フォンウィルブランド疾患3型；色素性乾皮症；地中海貧血症；レッシューナイラン症候群；プロテインC耐性；リソソーム疾患、例えばゴシェ病、ファブリー病；ムコ多糖症（MPS）1型（ハーレーーシャイエ症候群）、MPS II型（ハンター症候群）、MPS IIIA型（サンフィリオA症候群）、MPS IIIB型（サンフィリオB症候群）、MPS IIIC型（サンフィリオC症候群）、MPS IIID型（サンフィリオD症候群）、MPS IVA型（モルキオA症候群）、MPS IVB型（モルキオB症候群）、MPS VI型（マロトローラー症候群）、MPS VII型（スライ症候群）である。

#### 【0085】

好ましい態様において、標的DNAは突然変異を含み、そして選択されるDNA配列は突然変異を訂正することが可能な正常野生型配列を含む。

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は嚢胞性線維症膜貫通制御因子（CFTR）遺伝子中にある。好ましくは、突然変異は、CFTRタンパク質コード領域のコドン508のアミノ酸を改変するものであり、例えば突然変異は、CFTRタンパク質のコドン508のフェニルアラニンを除去する、3塩基対インフレーム欠失である。CFTRタンパク質におけるフェニルアラニン508のこの欠失は、嚢胞性線維症を有する被験者に高い割合で見られる。したがって、好ましい態様において、野生型CFTR遺伝子に見られるようなフェニルアラニン508をコードする配列を含む、選択されるDNA配列を用いて、突然変異CFTR遺伝子を標的とし、そして訂正することが可能である。

#### 【0086】



好ましい態様において、標的DNAは突然変異を含み、そして突然変異はヒトβ-グロビン遺伝子中にある。好ましくは、突然変異は、β-グロビン遺伝子の六番目のコドンのアミノ酸を改変するものであり、例えば突然変異は、β-グロビン遺伝子の六番目のコドンにおけるAからTの置換である。この突然変異は、β-グロビタンパク質において、鎌形赤血球貧血を有する被験者に見られる、グルタミン酸からバリンへの変化を導く。したがって、好ましい態様において、コドン6で野生型アミノ酸残基をコードする、選択されるDNA、例えば野生型β-グロビン遺伝子の六番目のコドン内で見られるように、Aを含む、選択されるDNA配列を用いて、突然変異β-グロビン遺伝子を標的とし、そして訂正することが可能である。

#### 【0087】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は因子VII遺伝子中にある。例えば、突然変異は、因子VII遺伝子のエクソン23、24、および／またはエクソン25にある可能性がある。好ましくは、突然変異は、因子VIIタンパク質コード領域のコード領域のコドン2209でアミノ酸を改変するものであり、例えば突然変異は、因子VIIのアミノ酸2209でアルギニンからグルタミンへの変化を導く、因子VII遺伝子のエクソン24におけるGからAの置換である。好ましくは、突然変異は、因子VIIタンパク質コード領域のコード領域のコドン2229でアミノ酸を改変するものであり、例えば突然変異は、因子VIIのアミノ酸2229でトリプトファンからシステインへの変化を導く、因子VII遺伝子のエクソン25におけるGからTの置換である。これらの突然変異は、中程度から重度の血友病Aと関連付けられてきている。したがって、好ましい態様において、因子VII遺伝子のコード領域のコドン2209で野生型アミノ酸をコードするDNA、または因子VII遺伝子のコード領域のコドン2229で野生型アミノ酸をコードするDNAいずれか、あるいは両方を含む、選択されるDNA配列を用いて、突然変異因子VII遺伝子を標的とし、そして訂正することが可能である。

#### 【0088】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は因子



I X遺伝子中にある。例えば、血友病Bを有する被験者において、突然変異の大部分は、因子I X遺伝子における点突然変異である。したがって、好ましい態様において、選択されるDNA配列は、血友病Bに関連する因子I X遺伝子中の1以上の点突然変異を標的とし、そして訂正するため、野生型因子I X遺伝子由来の少なくとも1つのヌクレオチドを有する1以上のヌクレオチドを含むことが可能である。

#### 【0089】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異はフォンウィルブランド因子遺伝子中にある。好ましくは、突然変異は、フォンウィルブランド遺伝子のエクソン18における位2679-2684の6シトシン伸長中の単一シトシン欠失である。この突然変異は、フォンウィルブランド疾患3型を有する被験者に、かなりの割合で見られる。フォンウィルブランド疾患3型に関連する他の突然変異、例えば点突然変異もまた、本明細書に記載されるように、改変可能である。したがって、好ましい態様において、野生型フォンウィルブランド遺伝子に見られる配列、例えばフォンウィルブランド遺伝子のエクソン18の位2679-2684の6つのシトシンを含む、選択されるDNA配列を用いて、突然変異フォンウィルブランド遺伝子を標的とし、そして訂正することが可能である。

#### 【0090】

好ましい態様において、標的DNAは突然変異を含み、そして突然変異は色素性乾皮症グループG (XP-G) 遺伝子中にある。好ましくは、突然変異は、XP-G遺伝子に見られる245塩基対エクソンの位19-21のアデニン伸長中の単一アデニン欠失である。この欠失は、色素性乾皮症を導く。したがって、好ましい態様において、XP-G遺伝子の野生型配列、例えばXP-G遺伝子の245塩基対エクソンの位19-21の3つのアデニンを含む、選択されるDNAを用いて、突然変異XP-G遺伝子を標的とし、そして訂正することが可能である。

#### 【0091】

別の好ましい態様において、改変は、選択されるDNA配列および標的DNA



、例えば染色体間の相同組換えを含む。

好ましい態様において、選択されるDNA配列は、1より多いヌクレオチドで標的DNAと異なり、例えば標的、または選択されるDNA配列が、非対領域、例えばループアウト領域を有するように、十分な数のヌクレオチドで、標的と異なる。こうした適用において、Msh2、Msh6、Msh3、Mlh1、Pms2、Mlh3、Pms1、またはそれらの組み合わせもまた、例えば複合体の一部として提供してもよい。

#### 【0092】

好ましい態様において、改変は、標的DNAへの選択される配列の組込みを含み、そして選択されるDNAは、標的上のあらかじめ選択される要素と、あらかじめ選択される関係にあるように組み込まれ、例えば一方が制御要素であり、そして他方がタンパク質をコードする配列である場合、制御要素が、タンパク質コード配列の発現を調節するよう機能する。選択される組込みを促進する隣接配列を用いてもよい。選択されるDNA配列は、選択される標的配列、例えば遺伝子またはコード配列の5'、3'、または該配列内に組み込むことが可能である。

#### 【0093】

好ましい態様において、改変は、選択されるDNA配列の組込みを含み、そして選択されるDNA配列は、制御配列、例えば外因性制御配列である。好ましい態様において、制御配列は、1以上の：プロモーター、エンハンサー、UAS、足場付着領域または転写因子結合部位を含む。好ましい態様において、制御配列は：メタロチオネイン-1遺伝子、例えばマウスメタロチオネイン遺伝子由来の制御配列、SV-40遺伝子由来の制御配列、サイトメガロウイルス遺伝子由来の制御配列、コラーゲン遺伝子由来の制御配列、アクチン遺伝子由来の制御配列、免疫グロブリン遺伝子由来の制御配列、HMG-CoAレダクターゼ遺伝子由来の制御配列、 $\gamma$ アクチン遺伝子由来の制御配列、転写活性化因子YY1遺伝子由来の制御配列、フィブロネクチン遺伝子由来の制御配列、またはEF-1 $\alpha$ 遺伝子由来の制御配列を含む。

#### 【0094】

好ましい態様において、選択されるDNA配列はエクソンを含む。好ましくは



、外因性エクソンは：CAP部位、ヌクレオチド配列ATG、および／または標的とされる内因性遺伝子とインフレーションのコードDNAを含む。

### 【0095】

好ましい態様において、選択されるDNA配列はスプライスドナー部位を含む。

好ましい態様において、選択されるDNA配列は、標的に組み込まれた際、内因性遺伝子の発現を制御するよう作用する、外因性制御配列を含む。選択されるDNAは、標的において、内因性遺伝子のコード領域の上流、または標的において、内因性遺伝子またはコード領域の内因性制御配列の上流に組み込むことが可能である。別の好ましい態様において、選択されるDNAは、内因性遺伝子またはコード領域の下流、あるいはイントロンまたは内因性遺伝子内に組み込むことが可能である。別の好ましい態様において、内因性遺伝子の内因性制御配列は不活性であり、例えば完全にまたは部分的に欠失される。

### 【0096】

好ましい態様において、選択されるDNA配列は内因性遺伝子上流であり、そして内因性遺伝子の第二のエクソンに連結される。

好ましい態様において、内因性遺伝子は：ホルモン、サイトカイン、抗原、抗体、酵素、凝固因子、輸送タンパク質、受容体、制御タンパク質、構造タンパク質または転写因子をコードする。好ましい態様において、内因性遺伝子は、以下のタンパク質：エリスロポエチン、カルシトニン、成長ホルモン、インスリン、インスリノトロピン、インスリン様増殖因子、副甲状腺ホルモン、 $\alpha 2$ -インターフェロン (IFNA2)、 $\beta$ -インターフェロン、 $\gamma$ -インターフェロン、神経増殖因子類、FSH $\beta$ 、TGF- $\beta$ 、腫瘍壊死因子、グルカゴン、骨増殖因子-2、骨増殖因子-7、TSH- $\beta$ 、インターロイキン1、インターロイキン2、インターロイキン3、インターロイキン6、インターロイキン11、インターロイキン12、CSF-顆粒球 (GCSF)、CSF-マクロファージ、CSF-顆粒球／マクロファージ、免疫グロブリン類、触媒性抗体類、プロテインキナーゼC、グルコセレブロシダーゼ、スーパーオキシドジスムターゼ、組織プラスミノゲン活性化因子、ウロキナーゼ、アンチトロンビンIII、DNアーゼ、



$\alpha$ -ガラクトシダーゼ、チロシンヒドロキシラーゼ、血液凝固因子V、血液凝固因子V I I、血液凝固因子V I I I、血液凝固因子I X、血液凝固因子X、血液凝固因子X I I I、アポリポタンパク質E、アポリポタンパク質A-I、グロビン類、低密度リポタンパク質受容体、I L-2受容体、I L-2アンタゴニスト類、 $\alpha$ -1-アンチトリプシン、免疫反応修飾剤類、 $\beta$ -グルコセラミダーゼ、 $\alpha$ -イズロニダーゼ、 $\alpha$ -L-イズロニダーゼ、グルコサミン-N-スルファターゼ、 $\alpha$ -N-アセチルグルコサミニダーゼ、アセチル補酵素A： $\alpha$ -グルコサミン-N-アセチルトランスフェラーゼ、N-アセチルグルコサミン-6-スルファターゼ、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、N-アセチルガラクトサミン-6-スルファターゼ、および可溶性C D 4のいずれかをコードする。

#### 【0097】

好ましい態様において、内因性遺伝子は卵胞刺激ホルモン $\beta$  (F S H  $\beta$ ) をコードし、そして選択されるD N A配列は、制御配列、例えばF S H  $\beta$  遺伝子の制御配列と配列が異なる制御配列を含む。好ましくは、選択されるD N A配列には、ターゲッティング配列が隣接し、例えばこうしたターゲッティング配列は、選択されるD N A配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲッティング配列は、F S H  $\beta$  コード領域(配列番号1)の5'の領域に相同である。好ましい態様において、ターゲッティング配列は、F S H  $\beta$  コード配列内で、またはF S H  $\beta$  コード配列上流で、相同組換えを指示する。好ましい態様において、ターゲッティング配列は、ヒトF S H  $\beta$  配列のヌクレオチド-7454から-1417(番号付けは翻訳開始部位に対する)に対応する配列番号2由来、またはヒトF S H  $\beta$  配列のヌクレオチド-696から-155に対応する配列番号3由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。

#### 【0098】

好ましい態様において、内因性遺伝子はインターフェロン $\alpha$  2 (I F N  $\alpha$  2) をコードし、そして選択されるD N A配列は、制御配列、例えばI F N  $\alpha$  2 遺伝子の制御配列と配列が異なる制御配列を含む。好ましくは、選択されるD N A配



列には、ターゲッティング配列が隣接し、例えばこうしたターゲッティング配列は、選択されるDNA配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲッティング配列は、IFN $\alpha$ 2コード領域の5'の領域に相同である。好ましい態様において、ターゲッティング配列は、IFN $\alpha$ 2コード配列上流の領域内で、相同組換えを指示する。好ましい態様において、ターゲッティング配列は、ヒトIFN $\alpha$ 2配列のヌクレオチドー4074からー511（番号付けは翻訳開始部位に対する）に対応する配列番号4由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。例えば、該配列は：ヒトIFN $\alpha$ 2配列のヌクレオチドー4074からー3796に対応する配列番号7由来の少なくとも20、30、50、または100ヌクレオチド；ヒトIFN $\alpha$ 2配列のヌクレオチドー582からー510に対応する配列番号8由来の少なくとも20、30、または50ヌクレオチド；ヒトIFN $\alpha$ 2配列のヌクレオチドー3795からー583に対応する配列番号9由来の少なくとも20、30、50、100、または1000ヌクレオチドを含んでもよい。

#### 【0099】

好ましい態様において、内因性遺伝子は顆粒球コロニー刺激因子（GCSF）をコードし、そして選択されるDNA配列は、制御配列、例えばGCSF遺伝子の制御配列と配列が異なる制御配列を含む。好ましくは、選択されるDNA配列には、ターゲッティング配列が隣接し、例えばこうしたターゲッティング配列は、選択されるDNA配列の1以上の端、好ましくは両端に存在する。好ましい態様において、ターゲッティング配列は、GCSFコード領域の5'の領域に相同である。好ましい態様において、ターゲッティング配列は、GCSFコード配列内で、またはGCSFコード配列上流で、相同組換えを指示する。好ましい態様において、ターゲッティング配列は、ヒトGCSF配列のヌクレオチドー6，578から101（番号付けは翻訳開始部位に対する）に対応する配列番号5由来の少なくとも20、30、50、100または1000の連続するヌクレオチドを含む。例えば、標的配列は、ヒトGCSF遺伝子のヌクレオチドー6，578からー364（番号付けは翻訳開始部位に対する）に対応する配列番号6由来の20、30、50、100または1000ヌクレオチドを含んでもよい。



## 【0100】

別の好ましい態様において、DNA配列はコード領域を含み、例えばDNA配列はタンパク質をコードする。好ましい態様において、コード領域は：ホルモン、サイトカイン、抗原、抗体、酵素、凝固因子、輸送タンパク質、受容体、制御タンパク質、構造タンパク質または転写因子をコードする。好ましい態様において、コード領域は、以下のタンパク質：エリスロポエチン、カルシトニン、成長ホルモン、インスリン、インスリノトロピン、インスリン様増殖因子、副甲状腺ホルモン、 $\beta$ -インターフェロン、 $\gamma$ -インターフェロン、神経増殖因子類、FSH $\beta$ 、TGF- $\beta$ 、腫瘍壊死因子、グルカゴン、骨増殖因子-2、骨増殖因子-7、TSH- $\beta$ 、インターロイキン1、インターロイキン2、インターロイキン3、インターロイキン6、インターロイキン11、インターロイキン12、CSF-顆粒球、CSF-マクロファージ、CSF-顆粒球/マクロファージ、免疫グロブリン類、触媒性抗体類、プロテインキナーゼC、グルコセレブロシダーゼ、スーパーオキシドジスムターゼ、組織プラスミノゲン活性化因子、ウロキナーゼ、アンチトロンビンIII、DNアーゼ、 $\alpha$ -ガラクトシダーゼ、チロシンヒドロキシラーゼ、血液凝固因子V、血液凝固因子VII、血液凝固因子VIII、血液凝固因子IX、血液凝固因子X、血液凝固因子XIII、アポリポタンパク質E、アポリポタンパク質A-I、グロビン類、低密度リポタンパク質受容体、IL-2受容体、IL-2アンタゴニスト類、 $\alpha$ -1-アンチトリプシン、免疫反応修飾剤類、 $\beta$ -グルコセラミダーゼ、 $\alpha$ -イズロニダーゼ、 $\alpha$ -L-イズロニダーゼ、グルコサミン-N-スルファターゼ、 $\alpha$ -N-アセチルグルコサミニダーゼ、アセチル補酵素A： $\alpha$ -グルコサミン-N-アセチルトランスフェラーゼ、N-アセチルグルコサミン-6-スルファターゼ、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、N-アセチルガラクトサミン-6-スルファターゼ、および可溶性CD4のいずれかをコードする。

## 【0101】

好ましい態様において、選択されるDNA配列は、内因性制御配列の下流または内因性遺伝子のコード領域の上流および遺伝子の内因性制御配列の下流の標的内に組み込むことが可能である。別の好ましい態様において、選択されるDNA



配列は、内因性遺伝子のコード領域が不活性である、例えば欠失されるように、内因性制御配列の下流に組み込むことが可能である。

#### 【0102】

別の側面において、本発明は、本明細書に記載される方法のいずれかによって作成される細胞を特徴とする。

別の側面において、本発明は、本明細書に記載される方法のいずれかによって、細胞において、遺伝子のタンパク質コード配列の発現を改変する方法を特徴とする。

#### 【0103】

好ましい態様において、該方法は、制御配列を含むDNA配列を有する、本明細書記載の複合体を、細胞内に導入し；標的とされるゲノム配列の改変を可能にして、相同組換え細胞を産生する条件下に、細胞を維持し；そして制御配列の調節下、遺伝子のタンパク質コード配列の発現を可能にする条件下で、相同組換え細胞を維持することを含む。

#### 【0104】

制御配列の調節下、遺伝子のタンパク質コード配列の発現を可能にする条件下で、相同組換え細胞を維持して、それにより、遺伝子のタンパク質コード配列の発現を改変する。

#### 【0105】

用語「相同」は、本明細書において、ターゲッティング配列および標的部位が、相同組換えを経ることが可能であるように、標的部位、例えば染色体DNA標的部位と同一であるか、または十分に類似であるターゲッティング配列を指す。相同組換えが有用な頻度で発生可能である限り、小さい割合の塩基対ミスマッチは許容可能である。

#### 【0106】

本明細書において、用語「野生型」は、疾患または機能障害に関連しない、例えば該疾患または機能障害を引き起こさず、それに貢献せず、それを左右せずまたは調節しない配列を指す。

#### 【0107】



本明細書において、「複合体」は、構成要素が共有または非共有結合によってカップリングされる、安定な会合を指す。

本発明の他の特徴および利点は、以下の詳細な説明から、そして請求項から明らかであろう。

### 発明の詳細な説明

#### 相同組換えを増進する剤

選択されるDNA配列および標的DNA、例えば染色体DNA間の相同組換えを促進するため、選択されるDNA配列と共に、相同組換えを増進する剤を提供することが可能である。相同組換えを増進する剤は、1以上の以下の機能：1) 選択されるDNA配列および組込みに選択される部位間の相同認識を増加させる機能；2) 選択されるDNA配列および組込みに選択される部位間の相同対形成を増加させる機能；3) 組換えDNA配列間の鎖侵入および鎖交換の効率を増加させる機能；4) 組換えの成熟産物への中間構造のプロセッシング効率を増加させる機能を有する。

#### 【0108】

相同組換えを増進する剤は、二本鎖DNA配列を含む混合物中で、細胞に導入可能であるし、DNA配列の投与直前または直後に導入可能であるし、またはDNA配列上に付着させ、例えばコーティングすることが可能である。相同組換えを増進する剤、例えばRad52、例えばhRad52、またはその断片で、全DNA配列をコーティングすることが可能であるし、あるいはDNA配列の1以上の端をコーティングすることが可能であり、例えばDNA配列の1以上の突出一本鎖端をコーティングすることが可能である。好ましくは、相同組換えを増進する剤は、DNA配列の突出一本鎖3'端または5'端の少なくとも部分をコーティングする。

#### 【0109】

相同組換えを増進する剤の例には：Rad52または機能するその断片；Rad51または機能するその断片；Rad54または機能するその断片；あるいはこれらのタンパク質またはこれらのタンパク質の断片の2以上の組み合わせが含まれる。相同組換えを増進する剤はまた、細胞内でも発現可能であるし、例えば



、上述の剤のいずれかをコードする核酸配列を細胞に導入することが可能である。

#### 【0110】

R a d 5 1断片が機能的であるかどうかの決定は、既知の技術によって、行うことが可能である。例えば、R a d 5 1断片の機能性は、例えばB a u m a n nら(1996) C e l l 87:757-766に記載されるように、当該技術分野に知られるi n v i t r oアッセイにおいて、相同対形成および鎖交換を仲介する能力に基づいて、決定可能である。簡潔には、h R a d 5 1をまず、環状s s D N Aとプレインキュベーションし、そしてその後、<sup>32</sup>P標識直線二重鎖D N Aを添加する。連結分子の形成および鎖交換の量は、電気泳動によって、決定可能である。さらに、R a d 5 1断片の機能性は、B e n s o nら(1994) E M B O J. 13:5764-5771に記載されるように、A T Pの存在下で、ニック形成された二重鎖に結合し、電子顕微鏡によって視覚化可能である、らせん核タンパク質フィラメントを形成する能力に基づいて、決定可能である。R a d 5 1の機能性はまた、機能するR a d 5 1タンパク質を欠く細胞において、D N A修復および相同組換えにおける欠陥を軽減する能力に基づいても、決定可能である。したがって、その非存在下に比較した際、上述のアッセイにおいて、陽性の影響を与えるのであれば、R a d 5 1断片が機能するかどうか、決定可能である。さらに、R a d 5 1断片によって与えられる陽性の影響の度合いは、全長R a d 5 1によって与えられる陽性の影響の度合いに比較可能である。

#### 【0111】

R a d 5 4断片の機能性は、当該技術分野に知られるアッセイ、例えばS w a g e m a k e r sら(1998) J. B i o l. C h e m. 273:28292-28297に記載されるアッセイにおいて、d s D N Aの存在下でA T Pを加水分解する能力に基づいて、決定可能である。さらに、R a d 5 4断片の機能性は、機能するR a d 5 4タンパク質を欠く細胞において、D N A修復および相同組換えにおける欠陥を軽減する能力に基づいて、決定可能である。

#### 【0112】

R a d 5 2および機能するその断片



標的DNAにおける選択される部位、例えば染色体DNAにおける選択される部位のDNA配列と共に提供されるRad52は、その非存在下で生じるであろうより、より高い率の部位の改変、例えば相同組換えを提供することが可能である。理論によって束縛されることは望ましくないが、Rad52は以下の機能：1)ヌクレアーゼ分解から全DNA配列を保護する機能；2)ヌクレアーゼ分解から、DNA配列の突出一本鎖端、例えば3'テールを保護する機能；3)DNA配列および組込みのため選択される部位間の相同認識を増加させる機能；並びに4)DNA配列および組込みのため選択される部位間の相同対形成を増加させる機能の1以上を提供することが可能であると考えられる。

#### 【0113】

Rad52は、Rad52の単離または遺伝子操作法によるコード配列の発現を含む、いくつかの方法で、得ることが可能である。例えばVan Dykeら(1999)Nature 398:728は、Sf9細胞からのhRad52の産生および精製を記載する。多様な種のRad52のヌクレオチド配列が知られる。例えばShenら(1995)Genomics 25(1):199-206(ネズミおよびヒトRad52)；Murisら(1994)Mutat. Res. 315(3):295-305(ネズミおよびヒトRad52)；Parkら(1995)J. Biol. Chem. 270(26):15467-15470(ヒトRad52)を参照されたい。

#### 【0114】

Rad52の断片は、例えばRad52またはその部分をコードする配列の発現によるか、あるいは遺伝子活性化による(好ましい方法)か、タンパク質分解的消化によるか、あるいは化学的合成による、いくつかの方法で、産生可能である。Rad52の内部または末端断片は、Rad52をコードする核酸の一端(末端断片のため)または両端(内部断片のため)由来の1以上のヌクレオチドを除去することによって、生成することが可能である。突然変異誘発DNAの発現は、Rad52ポリペプチド断片を生じる。したがって、「末端噛み取り(end-nibbling)」エンドヌクレアーゼまたは多様な制限酵素での消化は、Rad52断片のアレイをコードするDNAを生成することが可能である。R



a d 5 2 タンパク質の断片をコードするDNAはまた、ランダム剪断、制限消化または上に論じた方法の組み合わせによっても、生成することが可能である。

#### 【0115】

R a d 5 2 断片はまた、慣用的メリフィールド固相 f-M o c または t-B o c 化学反応などの、当該技術分野に知られる技術を用いて、化学的に合成することが可能である。例えば、R a d 5 2 ペプチドは、断片の重複を含まない望ましい長さの断片に任意に分割するか、または望ましい長さの重複断片に分割することが可能である。

#### 【0116】

R a d 5 2 断片が機能するかどうかの決定は、既知の技術によって、行うことが可能である。例えば、R a d 5 2 断片がヌクレアーゼ分解に対して保護することが可能であるかどうか決定するため、ヌクレアーゼ、例えばエクソヌクレアーゼまたはエンドヌクレアーゼの導入前に、末端標識直線化二本鎖DNA配列、例えば<sup>32</sup>P 標識直線化二本鎖DNA配列をR a d 5 2 断片とインキュベーションすることが可能である。その後、放出された標識、例えば<sup>32</sup>P の量を決定することが可能である。放出された標識の量は、R a d 5 2 断片がヌクレアーゼ分解に対して保護する能力の指標として役立つ。さらに、R a d 5 2 断片の機能性は、連結分子形成を刺激する能力に基づいて、決定可能である。R a d 5 2 断片の機能性は、Bensonら(1998) Nature 391:401-404に記載されるようなh R a d 5 1 駆動連結分子形成の刺激によって、in vitroで解析可能である。簡潔には、h R a d 5 1 をまず、環状ssDNAとプレインキュベーションし、そしてその後、<sup>32</sup>P 標識直線二重鎖DNAを添加する。連結分子の形成は、電気泳動によって決定可能である。R a d 5 2 の添加は、R a d 5 2 の非存在下の連結分子形成に比較した際、連結分子の形成を刺激する。したがって、非存在下の連結分子形成に比較した際、連結分子形成を刺激するならば、R a d 5 2 断片が機能するかどうか、決定可能である。さらに、R a d 5 2 断片による刺激の度合いは、全長R a d 5 2 刺激の度合いに比較可能である。さらに、R a d 5 2 断片の機能性は、Park(1995) J. Biol. Chem. 270:15467-15470に記載されるように、培養サル細胞



で過剰発現させた際、電離放射線に対する耐性を増加させ、そして相同組換えの率を増加させる能力に基づいて、決定可能である。

#### 【0117】

##### 非相同端連結を阻害する剤

非相同端連結を阻害する剤を用いて、非存在下で生じるであろうより、より高い率で、標的DNAにおける選択される部位に、DNA配列を提供することが可能である。非相同端連結は、二本鎖端間での不正確な融合を導く可能性があり、例えば再連結端は、挿入または欠失を有する可能性がある。非相同端連結を阻害する剤は、非相同端連結経路に関与する分子の発現および／または活性を阻害するいかなる剤であってもよい。例えば、Mre11、Rad50およびNbs1の複合体は、非相同端連結に関与する。したがって、例えばこの複合体の形成を阻害することによって、例えばこれらのタンパク質のいずれかに結合するか、またはこれらのタンパク質のいずれかの発現を阻害することによって、非相同端連結を阻害することが可能である。さらに、非相同端連結に関与する他のタンパク質には、Kuタンパク質、例えばKu70またはKu80、リガーゼ4 (Lig4) およびXrcc4が含まれる。

#### 【0118】

##### Ku不活性化剤

標的DNA中の選択される部位、例えば染色体DNA中の選択される部位で、DNA配列と共に、Ku不活性化剤を提供すると、その非存在下で生じるであろうより、より高い率の部位の改変、例えば相同組換えを提供することが可能である。Kuは、DNA不連続点に結合するおよそ70kDaおよび80kDaのヘテロ二量体であり、そして非相同端連結による二本鎖切断修復に役割を果たす。

「Ku80」はまた、「Ku86」とも称される可能性がある。

#### 【0119】

Ku不活性化剤は、Ku発現またはKu活性を阻害することが可能である。好ましくは、Ku不活性化剤は、KuまたはKuをコードするヌクレオチド配列と相互作用し、例えばそれと結合し、Ku発現またはKu活性を阻害する。好ましくは、Ku依存非相同端連結が阻害される。Ku阻害剤はKu70、Ku80ま



たは両方を阻害することが可能である。

#### 【0120】

K<sub>u</sub>を不活性化するのに使用可能な剤には、抗K<sub>u</sub>抗体およびK<sub>u</sub>結合分子、例えばK<sub>u</sub>に結合するランダム生成ペプチド、K<sub>u</sub>結合オリゴマーおよびポリマー、並びにアンチセンスK<sub>u</sub>核酸分子が含まれる。好ましくは、K<sub>u</sub>を不活性化する剤は、抗K<sub>u</sub>抗体およびK<sub>u</sub>結合分子などの局所投与可能な剤、例えばK<sub>u</sub>に結合するランダム生成ペプチド、およびK<sub>u</sub>結合オリゴマーまたはポリマーである。

#### 【0121】

好ましくは、K<sub>u</sub>不活性化剤は、K<sub>u</sub>と相互作用し、例えば結合する。K<sub>u</sub>タンパク質と相互作用する剤は、改変部位で、K<sub>u</sub>を局所的に不活性化することが可能である。

#### 【0122】

例えば、DNA配列および標的とされるDNAにごく近接して、K<sub>u</sub>不活性化剤を細胞に導入し、そしてそれにより、相同組換え部位で、局所的にK<sub>u</sub>を阻害する。K<sub>u</sub>不活性化剤は、二本鎖DNA配列を含む混合物中で、細胞に導入可能であるし、DNA配列の投与直前または直後に導入可能であるし、あるいはDNA配列またはDNA配列に関連するタンパク質、例えばRad52またはその断片に共有結合することが可能である。細胞はまた、抗K<sub>u</sub>抗体またはアンチセンスK<sub>u</sub>核酸分子などのK<sub>u</sub>不活性化剤とプレインキュベーションすることが可能である。

#### 【0123】

##### 抗K<sub>u</sub>抗体

抗K<sub>u</sub>抗体またはその断片を用いて、K<sub>u</sub>に結合させ、そしてそれによりK<sub>u</sub>活性を減少させることが可能である。抗K<sub>u</sub>抗体は、改変部位で、局所的にK<sub>u</sub>と相互作用するが、細胞において一般的にK<sub>u</sub>発現を阻害しないように、投与することが可能である。抗K<sub>u</sub>抗体には抗K<sub>u</sub>70および抗K<sub>u</sub>80抗体が含まれる。

#### 【0124】



K<sub>u</sub>タンパク質、またはその一部もしくは断片を免疫原として使用し、ポリクローナルおよびモノクローナル抗体調製のための標準的な技術を用いて、K<sub>u</sub>に結合する抗体を生成することが可能である。全長K<sub>u</sub>タンパク質が使用可能であるし、あるいはK<sub>u</sub>の抗原性ペプチド断片が免疫原として使用可能である。

#### 【0125】

典型的には、K<sub>u</sub>またはK<sub>u</sub>ペプチドを用いて、適切な被験者（例えばウサギ、ヤギ、マウスまたは他の哺乳動物）を免疫原で免疫することによって、抗体を調製する。適切な免疫原性調製は、例えば、K<sub>u</sub>をコードする配列の発現によって、または遺伝子活性化によって得られるK<sub>u</sub>タンパク質、あるいは化学的に合成されたK<sub>u</sub>ペプチドを含むことが可能である。例えば、明白に本明細書にその全体が援用される、米国特許第5,460,959号；および係属米国出願US SN 08/334,797；US SN 08/231,439；US SN 08/334,455；およびUS SN 08/928,881を参照されたい。K<sub>u</sub>のヌクレオチドおよびアミノ酸配列が知られ、そして例えばTakiguchiら(1996) Genomics 35(1):129-135に記載されている。調製は、フロイントの完全または不完全アジュバントなどのアジュバント、あるいは同様の免疫刺激性剤をさらに含むことが可能である。免疫原性K<sub>u</sub>調製を用いた適切な被験者の免疫は、ポリクローナル抗K<sub>u</sub>抗体反応を誘導する。

#### 【0126】

抗K<sub>u</sub>抗体またはその断片は、K<sub>u</sub>不活性化剤として使用可能である。抗K<sub>u</sub>抗体断片の例には、ペプシンなどの酵素で抗体を処理することによって生成可能な、F(v)、Fab、Fab' およびF(ab')<sub>2</sub>断片が含まれる。用語「モノクローナル抗体」または「モノクローナル抗体組成物」は、本明細書において、K<sub>u</sub>の特定のエピトープと免疫反応することが可能な抗原結合部位の1つの種のみを含む抗体分子集団を指す。したがってモノクローナル抗体組成物は、典型的には、免疫反応する特定のK<sub>u</sub>タンパク質に対する単一の結合親和性を示す。

#### 【0127】



さらに、標準的組換えDNA技術を用いて作成可能である、ヒトおよび非ヒト部分両方を含む、キメラおよびヒト化モノクローナル抗体などの、遺伝的操作法によって産生された抗Ku抗体が使用可能である。こうしたキメラおよびヒト化モノクローナル抗体は、当該技術分野に知られる標準的DNA技術、例えばRobinsonら、国際出願第PCT/US86/02269号；Akiraら、欧州特許出願184,187；Taniguchi, M.、欧州特許出願171,496；Morrisonら、欧州特許出願173,494号；Neubergerら、PCT国際公報第WO 86/01533号；Cabillyら、米国特許第4,816,567号；Cabillyら、欧州特許出願125,023；Betterら、Science 240:1041-1043, 1988；Liuら、PNAS 84:3439-3443, 1987；Liuら、J. Immunol. 139:3521-3526, 1987；Sunら、PNAS 84:214-218, 1987；Nishimuraら、Canc. Res. 47:999-1005, 1987；Woodら、Nature 314:446-449, 1985；およびShawら、J. Natl. Cancer Inst. 80:1553-1559, 1988；Morrison, S. L., Science 229:1202-1207, 1985；Oiら、BioTechniques 4:214, 1986；Winter、米国特許第5,225,539号；Jonesら、Nature 321:552-525, 1986；Verhoevenら、Science 239:1534, 1988；およびBeidlerら、J. Immunol. 141:4053-4060, 1988に記載される方法を用いた、遺伝子操作によって、産生可能である。

### 【0128】

さらに、標準的技術を用いて、Kuに対して向けられるヒトモノクローナル抗体が作成可能である。例えば、ヒトモノクローナル抗体は、トランスジェニックマウスまたは抗体産生ヒト細胞を移植した免疫不全マウスにおいて、生成可能である。こうしたマウスを生成する方法は、例えば、Woodら、PCT公報WO 91/00906、Kucheralapatiら、PCT公報WO 91/1



0741; Lonbergら、PCT公報WO 92/03918; Kayら、PCT公報WO 92/03917; Kayら、PCT公報WO 93/12227; Kayら、PCT公報94/25585; Rajewskyら、PCT公報WO 94/04667; Dittullioら、PCT公報WO 95/17085; Lonberg, N. ら(1994) *Nature* 368:856-859; Green, L. L. ら(1994) *Nature Genet.* 7:13-21; Morrison, S. L. ら(1994) *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Bruggemanら(1993) *Year Immunol* 7:33-40; Choiら(1993) *Nature Genet.* 4:117-123; Tuailonら(1993) *PNAS* 90:3720-3724; Bruggemanら(1991) *Eur J Immunol* 21:1323-1326; Duchosalら、PCT公報WO 93/05796; 米国特許第5, 411, 749号; McCuneら(1988) *Science* 241:1632-1639、Kamel-Reidら(1988) *Science* 242:1706; Spanopoulou(1994) *Genes & Development* 8:1030-1042; Shinkaiら(1992) *Cell* 68:855-868に記載される。ヒト抗体トランスジェニックマウスまたはヒト抗体産生細胞もしくは組織を移植した免疫不全マウスを、Kuまたは抗原性Kuペプチドで免疫し、そしてその後、これらの免疫マウス由来の脾臓細胞を用いて、ハイブリドーマを生成することが可能である。ハイブリドーマ産生法は公知である。

#### 【0129】

Kuに対するヒトモノクローナル抗体はまた、被験者のリンパ球由来のmRNAから調製される、免疫グロブリン軽鎖および重鎖cDNAを用いて、FabファージディスプレイライブラリーまたはscFvファージディスプレイライブラリーなどのコンビナトリアル免疫グロブリンライブラリーを構築することによっても、調製可能である。例えば、McCaffertyら、PCT公報WO 92/01047; Marksら(1991) *J. Mol. Biol.* 2



22: 581-597; および Griffiths ら (1993) EMBO J 12: 725-734 を参照されたい。さらに、既知のヒト抗体を突然変異させることによって、抗体可変領域のコンビナトリアルライブラリーが生成可能である。例えば、Ku に結合することが知られるヒト抗体の可変領域を、例えばランダムに改変された突然変異誘発オリゴヌクレオチドを用いて、突然変異可変領域のライブラリーを生成することによって、突然変異させ、これをその後、Ku に結合するかスクリーニングすることが可能である。免疫グロブリン重鎖および／または軽鎖の CDR 領域内にランダム突然変異誘発を誘導する方法、ランダム化した重鎖および軽鎖を組み合わせて、対形成およびスクリーニング法を形成する方法は、例えば Barbass ら、PCT 公報 WO 96/07754; Barbass ら (1992) Proc. Nat' l Acad. Sci. USA 89: 4457-4461 に見出すことが可能である。

#### 【0130】

免疫グロブリンライブラリーは、好ましくは繊維状ファージ由来の、ディスプレイパッケージ集団によって発現させ、抗体ディスプレイライブラリーを形成することが可能である。抗体ディスプレイライブラリーを生成するのに特に使用しやすい方法および試薬の例は、例えば、Ladner ら、米国特許第 5, 223, 409 号; Kang ら、PCT 公報 WO 92/18619; Dower ら、PCT 公報 WO 91/17271; Winter ら、PCT 公報 WO 92/20791; Markland ら、PCT 公報 WO 92/15679; Breitling ら、PCT 公報 WO 93/01288; McCafferty ら、PCT 公報 WO 92/01047; Garrard ら、PCT 公報 WO 92/09690; Ladner ら、PCT 公報 WO 90/02809; Fuchs ら (1991) Bio/Technology 9: 1370-1372; Hay ら (1992) Hum Antibod Hybridomas 3: 81-85; Huse ら (1989) Science 246: 1275-1281; Griffiths ら (1993) 上記; Hawkins ら (1992) J Mol Biol 226: 889-896; Clackson ら (1991) Nature 352: 624-628; Gram ら (1992)



PNAS 89: 3576-3580; Garradら (1991) Bio/Technology 9: 1373-1377; Hoogenboomら (1991) Nuc Acid Res 19: 4133-4137; および Barrbasら (1991) PNAS 88: 7978-7982に見出すことが可能である。ディスプレイパッケージ (例えば繊維状ファージ) 表面上にディスプレイしたら、抗体ライブラリーをスクリーニングし、Kuに結合する抗体を発現するパッケージを同定し、そして単離する。好ましい態様において、ライブラリーの一次スクリーニングは、固定Kuを用いてパニングすることを伴い、そして固定Kuに結合する抗体を発現するディスプレイパッケージを選択する。

### 【0131】

Kuに対するモノクローナル抗体はまた、例えばNeomarkers (カリフォルニア州フレモント) からも商業的に入手可能である。

#### Ku結合分子

Ku結合ペプチド、例えばランダム生成ペプチド、およびKu結合オリゴマーまたはポリマーなどのKuに結合する分子は、Ku不活性化剤として、使用可能である。こうした分子は、Kuタンパク質に結合し、そして非相同端連結などのKuの少なくとも1つの活性を阻害することが可能である。

### 【0132】

Ku結合オリゴマーの例は、その内容が本明細書に援用されるWO 99/33971に示される。こうしたオリゴマーは、ヌクレオチド、ヌクレオチド類似体 (analog)、または組み合わせで構成可能である。好ましくは、オリゴマーはリボヌクレオチドで構成される。これらのKuオリゴマーを用いて、Kuに結合させるか、またはKuと相互作用するタンパク質を同定することが可能である。これらのオリゴマーを用いてKu結合ペプチドを同定する方法は、WO 99/33971に記載される。

### 【0133】

さらに、ランダム生成ペプチドを、Kuに結合する能力に関してスクリーニングすることが可能である。例えば、生成された突然変異遺伝子産物をスクリーニングする多様な技術が当該技術分野に知られる。大きな遺伝子ライブラリーをス



クリーニングする技術は、しばしば、遺伝子ライブラリーを複製可能発現ベクター内にクローニングし、生じたベクターライブラリーで、適切な細胞を形質転換し、そして望ましい活性、例えばKuへの結合の検出が、その産物を検出する遺伝子をコードするベクターの、比較的容易な単離を促進する条件下で、遺伝子を発現させることを含む。以下に記載される技術の各々は、例えばランダム突然変異誘発技術によって、生成された多数の配列をスクリーニングするための高処理解析に受け入れられる。

### 【0134】

#### ディスプレイライブラリー

Ku結合ペプチドをスクリーニングする別のアプローチにおいて、候補ペプチドを細胞またはウイルス粒子表面上にディスプレイし、そして特定の細胞またはウイルス粒子が、ディスプレイされた産物を介してKuタンパク質に結合する能力を、「パニングアッセイ」で検出する。例えば、遺伝子ライブラリーを、細菌細胞の表面膜タンパク質の遺伝子内にクローニングし、そして生じた融合タンパク質をパニングによって検出することが可能である(Ladnerら、WO 88/06630; Fuchsら(1991) Bio/Technology 9:1370-1371; およびGowardら(1992) TIBS 18:136-140)。同様の方式で、検出可能標識リガンドを用いて、潜在的に機能するペプチド相同体に関してスコア付けが可能である。蛍光標識リガンドを用いて、リガンド結合活性を保持する相同体を検出可能である。蛍光標識リガンドを使用すると、蛍光顕微鏡下で、細胞を視覚的に検査し、そして分離することが可能になるか、または細胞の形態が許す場合、蛍光活性化細胞分取によって、分離することが可能になる。

### 【0135】

遺伝子ライブラリーは、ウイルス粒子表面上の融合タンパク質として発現させることが可能である。例えば、繊維状ファージ系において、異質の(foreign)ペプチド配列を感染性ファージの表面上に発現させ、それにより2つの有意な利点を与えることが可能である。まず、これらのファージは、1ミリリットルあたり、 $10^{13}$  ファージをはるかに越える濃度で、アフィニティマトリック



スに適用可能であるため、多数のファージを一度にスクリーニングすることが可能である。第二に、各感染性ファージは、その表面上に遺伝子産物をディスプレイするため、特定のファージが低収量でアフィニティマトリックスから回収されたら、別の感染周期によって、そのファージを増幅することが可能である。ほぼ同一の大腸菌 (*E. coli*) 繊維状ファージ、M13、fd、および f1 の群が、ファージディスプレイライブラリーにおいて、最も頻繁に用いられる。ファージ gIII または gVIII コートタンパク質いずれかを用いて、ウイルス粒子の最終的なパッケージングを乱すことなく、融合タンパク質を生成することが可能である。異質のエピトープを pIII の NH<sub>2</sub> 末端で発現させ、そしてこうしたエピトープを持つファージを、このエピトープを欠く多量の過剰なファージから回収することが可能である (Ladner ら、PCT 公報 WO 90/02909; Garrard ら、PCT 公報 WO 92/09690; Marks ら (1992) *J. Biol. Chem.* 267:16007-16010; Griffiths ら (1993) *EMBO J* 12:725-734; Clackson ら (1991) *Nature* 352:624-628; および Barbas ら (1992) *PNAS* 89:4457-4461)。

### 【0136】

一般的なアプローチは、大腸菌のマルトース受容体 (外膜タンパク質、LamB) を、ペプチド融合パートナーとして用いる (Charbit ら (1986)

*EMBO* 5, 3029-3037)。LamB 遺伝子をコードするプラスミドにオリゴヌクレオチドを挿入し、タンパク質の細胞外ループの 1 つに融合させたペプチドを産生しておく。これらのペプチドは、リガンド、例えば抗体への結合に利用可能であり、そして細胞を動物に投与する際、免疫反応を引き出すことが可能である。他の細胞表面タンパク質、例えば OmpA (Schorr ら (1991) *Vaccines* 91, pp. 387-392), PhoE (Agterberg ら (1990) *Gene* 88, 37-45)、および PAL (Fuchs ら (1991) *Bio/Tech* 9, 1369-1372) と共に、巨大細菌表面構造が、ペプチドディスプレイ用のビヒクルとして



、利用されてきている。重合して、遺伝子情報の細菌間交換のための纖毛―導管を形成するタンパク質であるピリンに、ペプチドを融合させてもよい (Thiryら (1989) Appl. Environ. Microbiol. 55, 984-993)。他の細胞と相互作用する際の役割のため、纖毛は、細胞外環境へのペプチドの提示に有用な支持体を提供する。ペプチドディスプレイに用いられる別の巨大表面構造は、細菌輸送器官、鞭毛である。サブユニットタンパク質、フラジェリンへのペプチドの融合は、宿主細胞上に多くのペプチドコピーの高密度アレイを提供する (Kuwayamaら (1988) Bio/Tech. 6, 1080-1083)。他の細菌種の表面タンパク質もまた、ペプチド融合パートナーとして利用されてきている。例には、スタフィロコッカス属 (Staphylococcus) プロテインAおよびナイセリア属 (Neisseria) の外膜プロテアーゼ IgA が含まれる (Hanssonら (1992) J. Bacteriol. 174, 4239-4245 および Klausnerら (1990) EMBO J. 9, 1991-1999)。

### 【0137】

上述の繊維状ファージ系および LamB 系において、ペプチドおよびそのコード DNA 間の物理的連結は、その表面上にペプチドを所持する粒子 (細胞またはファージ) 内に DNA を包含することによって起こる。ペプチド捕捉は、粒子および内部の DNA を捕捉する。別の計画は、DNA 結合タンパク質、LacI を用いて、ペプチドおよび DNA 間の連結を形成する (Cullら (1992) PNAS USA 89:1865-1869)。この系は、3' 端にオリゴヌクレオチドクロニング部位を持つ LacI 遺伝子を含むプラスミドを用いる。アラビノースによる調節誘導下、LacI ペプチド融合タンパク質を産生する。この融合は、LacI が LacO オペレーター (LacO) として知られる短い DNA 配列に結合する天然の能力を保持する。発現プラスミド上に 2 コピーの LacO を備えることによって、LacI-ペプチド融合体は、これをコードするプラスミドに緊密に結合する。各細胞内のプラスミドは、単一のオリゴヌクレオチド配列のみを含み、そして各細胞は、単一のペプチド配列のみを発現するため



、ペプチドは、その合成を指示するDNA配列と特異的にそして安定に会合する。ライブラリーの細胞を穏やかに溶解し、そしてペプチド-DNA複合体を固定受容体のマトリックスに曝露し、活性ペプチドを含む複合体を回収する。その後、会合プラスミドDNAを、増幅のため、再度、細胞に導入し、そしてDNAを配列決定して、ペプチドリガンドの同一性を決定する。方法の実際的な有用性の立証として、ドデカペプチドの大きなランダムライブラリーを作成し、そしてオピオイドペプチド、ダイノルフィンBに対して作成されたモノクローナル抗体上で選択した。すべてダイノルフィンBの6残基部分に対応するコンセンサス配列によって関連する、ペプチドのコホートを回収した (Cullら (1992) Proc. Natl. Acad. Sci. U. S. A. 89-1869)。

#### 【0138】

この計画は、時に、ペプチド-オン-プラスミドと称され、ファージディスプレイ法とは、2つの重要な点で異なる。まず、ペプチドは、融合タンパク質のC末端に付着し、未結合 (free) カルボキシ末端を有するペプチドとして、ライブラリーメンバーのディスプレイを生じる。繊維状ファージコートタンパク質、pIIIおよびpVIIはどちらも、そのC末端を通じてファージに係留され；そして外側に伸長するN末端ドメイン内に、ゲストペプチドが置かれる。いくつかの設計では、ファージにディスプレイされるペプチドは、融合タンパク質のちょうどアミノ末端で提示される (Cwirlaら (1990) Proc.

Natl. Acad. Sci. U. S. A. 87, 6378-6382)。第二の相違は、ライブラリーに実際存在するペプチド集団に影響を与える、一組の生物学的バイアスである。LacI融合分子は、宿主細胞の細胞質に制限される。ファージコート融合体は、翻訳中に短時間、細胞質に曝露されるが、内膜を通じて周辺質区画に迅速に分泌され、C末端疎水性ドメインによって膜に係留されたまま、ペプチドを含むN末端は、ファージ粒子への組み立てを待つ間、周辺質に突出している。LacIおよびファージライブラリー中のペプチドは、異なるタンパク質分解活性への曝露の結果、有意に異なる可能性がある。ファージコートタンパク質は、ファージへの取り込みの序幕として、内膜を渡る輸



送およびシグナルペプチダーゼプロセッシングを必要とする。特定のペプチドは、これらの過程に有害な影響を発揮し、そしてライブラリー中で過少提示される（Gallopら（1994） J. Med. Chem. 37（9）：1233-1251）。これらの特定のバイアスは、LacIディスプレイ系の要因ではない。

### 【0139】

組換えランダムライブラリーに利用可能な小ペプチドの数は莫大である。 $10^7$  -  $10^9$  の独立クローンのライブラリーが日常的に調製される。 $10^{11}$  組換え体と同程度の大きさのライブラリーが生成されてきているが、この大きさは、クローンライブラリーの実際的な限界に近づいている。ライブラリーサイズのこの限界は、ランダム化部分を含むDNAを宿主細菌細胞に形質転換する工程で起こる。この限界を回避するため、ポリソーム複合体中の新生ペプチドのディスプレイに基づく *in vitro* 系が最近、開発されてきている。このディスプレイライブラリー法は、現在利用可能なファージ／ファージミドまたはプラスミドライブラリーより、3-6桁大きいライブラリーを産生する可能性を有する。さらに、ライブラリーの構築、ペプチドの発現、およびスクリーニングは、完全に細胞不含形式で行う。

### 【0140】

この方法の1つの適用において（Gallopら（1994） J. Med. Chem. 37（9）：1233-1251）、 $10^{12}$  デカペプチドをコードする分子DNAライブラリーを構築し、そして大腸菌S30で発現したライブラリーを、転写／翻訳系と *in vitro* カップリングした。リボソームをmRNA上に引き止めるように条件を選択し、かなりの割合のRNAをポリソーム中に集積させ、そしてコードRNAにまだ連結している新生ペプチドを含む複合体を生じた。ポリソームは、より慣用的な組換えペプチドディスプレイライブラリーをスクリーニングするのと同ほとんど同じ方式で、固定受容体上でアフィニティー精製するのに十分に丈夫である。結合複合体由来のRNAを回収し、cDNAに変換し、そしてPCRによって増幅して、合成およびスクリーニングの次の周期のテンプレートを生じる。ポリソームディスプレイ法をファージディスプ



レー系とカップリングしてもよい。数周期のスクリーニング後、ポリソームの濃縮プール由来のcDNAを、ファージミドベクターにクローニングした。このベクターは、コートタンパク質に融合したペプチドをディスプレイするペプチド発現ベクターとして、そしてまたペプチド同定のためのDNA配列決定ベクターとして、役立つ。ファージ上でポリソーム由来ペプチドを発現させることによって、この形式の親和性選択法を続けることが可能であるし、あるいはファージELISAにおける結合活性に関して、または完成ファージELISA (Barrettら (1992) Anal. Biochem 204, 357-364) における結合特異性に関して、個々のクローン上のペプチドをアッセイすることが可能である。活性ペプチドの配列を同定するため、ファージミド宿主によって産生されたDNAを配列決定する。

#### 【0141】

##### アンチセンスKu核酸配列

Kuをコードするヌクレオチドに対してアンチセンスである核酸分子は、Ku発現を阻害する不活性化剤として使用可能である。「アンチセンス」核酸には、Kuをコードする「センス」核酸に相補的である、例えば二本鎖cDNA分子のコード鎖に相補的であるかまたはmRNA配列に相補的である、ヌクレオチド配列が含まれる。したがって、アンチセンス核酸は、センス核酸と水素結合を形成することが可能である。アンチセンス核酸は、全Kuコード鎖に相補的であってもよいし、またはその部分だけに相補的であってもよい。例えば、Kuをコードするヌクレオチド配列のコード鎖の「コード領域」にアンチセンスであるアンチセンス核酸分子が使用可能である。

#### 【0142】

Kuをコードするコード鎖配列が、例えば、Takiguchiら (1996) Genomics 35 (1) : 129-135およびGenbank寄託番号L35932に開示されることを考慮し、ワトソンおよびクリック塩基対形成の規則にしたがって、アンチセンス核酸を設計することが可能である。アンチセンス核酸分子は、Ku mRNAの全コード領域に相補的であってもよいが、より好ましくは、Ku mRNAのコードまたは非コード領域の部分に対しての



みアンチセンスであるオリゴヌクレオチドである。例えば、アンチセンスオリゴヌクレオチドは、Ku mRNAの翻訳開始部位を取り巻く領域に相補的であってもよい。アンチセンスオリゴヌクレオチドは、例えば、長さ約5、10、15、20、25、30、35、40、45または50ヌクレオチドであってもよい。アンチセンス核酸は、当該技術分野に知られる方法を用いた、化学合成および酵素連結を用いて、構築可能である。例えば、アンチセンス核酸（例えばアンチセンスオリゴヌクレオチド）は、天然存在ヌクレオチドを用いて、あるいは分子の生物学的安定性を増加させるか、またはアンチセンスおよびセンス核酸間に形成される二重鎖の物理的安定性を増加させるように設計した、多様な修飾ヌクレオチドを用いて、化学的に合成可能であり、例えばホスホロチオエート誘導体およびアクリジン置換ヌクレオチドが使用可能である。アンチセンス核酸を生成するのに使用可能な修飾ヌクレオチドの例には、5-フルオロウラシル、5-ブロモウラシル、5-クロロウラシル、5-ヨードウラシル、ヒポキサンチン、キサンチン、4-アセチルシトシン、5-（カルボキシヒドロキシメチル）ウラシル、5-カルボキシメチルアミノメチル-2-チオウリジン、5-カルボキシメチルアミノメチルウラシル、ジヒドロウラシル、ベータ-D-ガラクトシルケオシン、イノシン、N6-イソペンテニルアデニン、1-メチルグアニン、1-メチルイノシン、2, 2-ジメチルグアニン、2-メチルアデニン、2-メチルグアニン、3-メチルシトシン、5-メチルシトシン、N6-アデニン、7-メチルグアニン、5-メチルアミノメチルウラシル、5-メトキシアミノメチル-2-チオウラシル、ベータ-D-マンノシルケオシン、5'-メトキシカルボキシメチルウラシル、5-メトキシウラシル、2-メチルチオ-N6-イソペンテニルアデニン、ウラシル-5-オキシ酢酸(v)、ワイブトキソシン、プソイドウラシル、ケオシン、2-チオシトシン、5-メチル-2-チオウラシル、2-チオウラシル、4-チオウラシル、5-メチルウラシル、ウラシル-5-オキシ酢酸メチルエステル、ウラシル-5-オキシ酢酸(v)、5-メチル-2-チオウラシル、3-（3-アミノ-3-N-2-カルボキシプロピル）ウラシル、(acp3)w、および2, 6-ジアミノプリンが含まれる。あるいは、アンチセンス核酸は、核酸がアンチセンス方向にサブクローニングされている発現ベクターを



用いて、生物学的に産生可能である（すなわち、挿入核酸から転写されたRNAは、目的の標的核酸に対してアンチセンス方向のものである）。

### 【0143】

#### 外因性DNA配列

細胞に提供しようとする、例えば導入しようとするDNA配列は、細胞において、標的配列を改変させることが可能である。例えば、10、8、5、4、3、2未満、または1ヌクレオチド、例えば置換、欠失、または挿入により、標的DNAと異なる、選択されるDNA配列を導入することが可能である。選択されるDNA配列はまた、1より多いヌクレオチド、標的配列と異なってもよく、例えば選択されるDNA配列が、非対領域、例えばループアウト領域を有するよう、いくつかのヌクレオチド、標的配列と異なる。これらの改変は、標的配列発現を修飾することが可能である。修飾される配列発現は：細胞において、通常サイレントである（発現されない）配列、例えばコードDNA配列、例えば細胞に通常見られるコード配列を活性化し；細胞において、正常レベルより低く発現される配列、例えばコードDNA配列、例えば細胞に通常見られるコード配列の発現を増加させ；細胞において、通常、欠損型で発現される配列、例えばコードDNA配列、例えば細胞に通常見られるコード配列を発現させ；細胞の通常のパターンと異なるように、配列、例えばコードDNA配列、例えば細胞に通常見られるコード配列の制御または誘導パターンを変化させ；配列、例えばコードDNA配列、例えば細胞に通常見られるコード配列の発現を、細胞における通常の発現レベルより減少させることを含む。

### 【0144】

10、8、5、4、3、2未満、または1ヌクレオチド、例えば置換、欠失、または挿入により、標的DNAと異なる、選択されるDNA配列を導入することが可能である。例えば、標的とされる配列は、野生型配列と、10、8、5、4、3、2未満、または1ヌクレオチド、異なる可能性がある。好ましくは、標的とされる配列は、点突然変異、例えば挿入、欠失または置換から生じる突然変異により、野生型配列と異なる。好ましくは、標的配列、例えば遺伝子中の突然変異は、疾患または機能障害に関連し、例えば該疾患または機能障害を調節する。



突然変異、例えば点突然変異が、疾患または機能障害に関連付けられている遺伝子の例には、限定されるわけではないが、嚢胞性線維症膜貫通制御因子（C F T R）遺伝子、 $\beta$ -グロビン遺伝子、因子V I I I 遺伝子、因子I X 遺伝子、フォンウィルブランド因子遺伝子、色素性乾皮症グループG（X P - G）遺伝子が含まれる。標的配列を改変させるため、選択されるD N A 配列は、突然変異を訂正することが可能な正常野生型配列を含んでもよい。本明細書に記載される方法にしたがって改変可能な、いくつかの遺伝子障害および遺伝子がある。

#### 【0145】

別の側面において、選択されるD N A 配列はまた、1より多いヌクレオチド、標的配列と異なってもよく、例えば選択されるD N A 配列が、非対領域、例えばループアウト領域を有するよう、いくつかのヌクレオチド、標的配列と異なる。例えば、選択されるD N A 配列は、標的のあらかじめ選択される要素と相同組換えされることが可能であり、例えば一方が制御要素であり、そして他方がタンパク質をコードする配列である場合、制御要素が、タンパク質コード配列の発現を制御する。選択されるD N A 配列は、制御配列、例えば外因性制御配列であってもよい。制御配列は、プロモーター、エンハンサー、U A S、足場付着領域および転写因子結合部位を含む。さらに、選択されるD N A 配列はまた、エクソン、イントロン、C A P 部位、ヌクレオチド配列A T G、マーカー、例えば選択マーカー、スプライスドナー部位および／または標的配列とインフレームのコードD N A も含むことが可能である。選択されるD N A 配列はまた、コード領域、例えばタンパク質をコードするD N A 配列も含むことが可能である。

#### 【0146】

コード配列が内因性であってもよく、例えば選択されるD N A 配列は制御配列であるか、または選択されるD N A 配列がコード領域を含んでもよく、すなわちコード領域は外因性である。コード領域は、多様なタンパク質をコードしていてもよい。こうしたタンパク質の例には：エリスロポエチン、カルシトニン、成長ホルモン、インスリン、インスリノトロピン、インスリン様増殖因子、副甲状腺ホルモン、 $\alpha$ 2-インターフェロン（I F N A 2）、 $\beta$ -インターフェロン、 $\gamma$ -インターフェロン、神経増殖因子類、F S H  $\beta$ 、T G F -  $\beta$ 、腫瘍壊死因子、



グルカゴン、骨増殖因子-2、骨増殖因子-7、TSH- $\beta$ 、インターロイキン1、インターロイキン2、インターロイキン3、インターロイキン6、インターロイキン11、インターロイキン12、CSF-顆粒球(GCSF)、CSF-マクロファージ、CSF-顆粒球/マクロファージ、免疫グロブリン類、触媒性抗体類、プロテインキナーゼC、グルコセレブロシダーゼ、スーパーオキシドジスムターゼ、組織プラスミノーゲン活性化因子、ウロキナーゼ、アンチトロンビンIII、DNアーゼ、 $\alpha$ -ガラクトシダーゼ、チロシンヒドロキシラーゼ、血液凝固因子V、血液凝固因子VII、血液凝固因子VIII、血液凝固因子IX、血液凝固因子X、血液凝固因子XIII、アポリポタンパク質E、アポリポタンパク質A-I、グロビン類、低密度リポタンパク質受容体、IL-2受容体、IL-2アンタゴニスト類、 $\alpha$ -1-アンチトリプシン、免疫反応修飾剤類、 $\beta$ -グルコセラミダーゼ、 $\alpha$ -イズロニダーゼ、 $\alpha$ -L-イズロニダーゼ、グルコサミン-N-スルファターゼ、 $\alpha$ -N-アセチルグルコサミニダーゼ、アセチル補酵素A： $\alpha$ -グルコサミン-N-アセチルトランスフェラーゼ、N-アセチルグルコサミン-6-スルファターゼ、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、N-アセチルガラクトサミン-6-スルファターゼ、および可溶性CD4が含まれる。これらのタンパク質をコードする配列は既知である。

#### 【0147】

用語、外因性は、本明細書に記載される方法によって、細胞に導入される配列を指す。外因性配列は、細胞に存在する内因性配列と同一の配列または異なる配列を有することが可能である。

#### 【0148】

好ましくは、DNA配列は直鎖配列である。

#### ターゲッティング配列または配列類

ターゲッティング配列または配列類は、標的とされる配列、例えば標的とされる遺伝子を含む細胞のゲノム内への相同組換えを可能にするDNA配列である。用語「ターゲッティング配列」および「隣接配列」は、本明細書において、交換可能に用いられる。ターゲッティング配列は、一般的に、得られるような細胞ゲノムに通常存在するDNA配列に相同である（すなわち細胞DNAに同一である



か、またはターゲッティング配列および細胞DNAが相同組換えを経ることが可能であるように十分に類似である) DNA配列である。例えば、ターゲッティング配列は：コードまたは非コードDNA、目的の遺伝子の転写開始部位上流にある配列、目的の遺伝子内にある配列、または目的の遺伝子の転写停止部位の下流にある配列、あるいは先の修飾を通じてゲノムに存在する配列に、十分に相同であってもよい。用いるターゲッティング配列または配列類は、選択されるDNA配列を挿入しようとする部位または標的とされる配列を改変しようとする部位に照合して選択する。

#### 【0149】

1以上のターゲッティング配列が使用可能である。好ましくは、選択されるDNA配列には、2つのターゲッティング配列が隣接する。ターゲッティング配列は、遺伝子またはコード配列内に(エクソンおよび/またはイントロンの配列など)、遺伝子のコード配列にすぐ隣接して(例えばターゲッティング配列および遺伝子のコード領域間に10、5、4、3、2、1未満のヌクレオチドを含み、またはさらなるヌクレオチドをまったく含まず)、遺伝子のコード配列上流に(上流非コード領域の配列または内因性プロモーター配列など)、または遺伝子のコード配列の上流でそして少し離れて(内因性プロモーターの上流の配列など)存在してもよい。ターゲッティング配列または配列類は、現在知られるかまたは配列決定された、標的とされる配列の領域、および/または構造的に性質決定されていないが、当業者によって、制限酵素を用いてマッピングし、そして決定することが可能な、さらに上流の領域を含むことが可能である。

#### 【0150】

ターゲッティング配列を用いて、内因性遺伝子のコード配列にすぐ隣接して、その上流に、またはかなり離れて、制御配列を含むDNA配列を挿入することが可能である。あるいはまたはさらに、産生されるRNAまたはタンパク質の構造または安定性に影響を与える配列を、ターゲッティングによって、置換するか、除去するか、付加するか、または別の方式で修飾することが可能である。例えば、RNA分子のRNA安定性要素、スプライス部位、および/またはリーダー配列を修飾し、RNA分子の機能、安定性、および/または翻訳可能性を改善する



かまたは改変することが可能である。タンパク質の輸送、分泌、または機能特性を増進するかまたは修飾するため、シグナル配列、プロペプチド配列、活性部位、および／または構造配列などのタンパク質配列もまた、改変することが可能である。タンパク質配列はまた、突然変異、例えば点突然変異を含むタンパク質をコードする遺伝子における部位を標的とすることによっても、改変、例えば訂正可能である。

### 【0151】

1つの側面において、ターゲッティング配列は、ヒト卵胞刺激ホルモン $\beta$  (FSH $\beta$ )の部分に相同であってもよい。FSHは、正常な生殖生理において、卵母細胞および精子の維持および発生に必須の役割を果たす性腺刺激ホルモンである。FSHは2つのサブユニット、 $\alpha$ および $\beta$ を含み、後者は、FSHの生物学的特異性に関与している。既定のターゲッティング配列が相同である標的部位は、FSH $\beta$ 遺伝子のエクソンおよび／またはイントロン内に、FSH $\beta$ コード領域の上流にそしてすぐ隣接して、またはFSH $\beta$ コード領域の上流にそして少し離れて存在していてもよい。例えば、2つのターゲッティング配列の第一のもの（または構築物中に1つのターゲッティング配列しかない場合、全ターゲッティング配列）は、FSH $\beta$ コード配列の上流のゲノム領域由来であってもよい。例えば、このターゲッティング配列は、配列番号1の部分、例えば位-7, 454から-1, 417（配列番号2）または位-696から-155（配列番号3）に対応する配列由来の少なくとも20、30、50、100、または1000の連続するヌクレオチドを含んでもよい。2つのターゲッティング配列の第二のものは、コード配列上流のゲノム領域を標的としてもよいし（例えば、やはり配列番号2または3の部分を含んでもよい）、あるいは遺伝子のエクソンまたはイントロンを標的としてもよい。FSH $\beta$ を標的とするのに使用可能な配列は、その内容が完全に本明細書に援用される、米国特許出願第09/305, 639号にさらに記載される。

### 【0152】

ターゲッティング配列は、ヒトインターフェロン $\alpha 2$  (IFN $\alpha 2$ )の部分に相同であってもよい。インターフェロン $\alpha$ は、染色体9の短腕にクラスター



を形成する14の遺伝子を含む複雑な遺伝子ファミリーを構成する。IFN $\alpha$ 2遺伝子を含む、これらの遺伝子のいずれも、イントロンを持たない。インターフェロン $\alpha$ は、マクロファージ、T細胞およびB細胞と共に、多くの他の細胞種によって産生される。インターフェロン $\alpha$ は、かなりの抗ウイルス効果を有し、そしてパピローマウイルス、B型およびC型肝炎ウイルス、ワクシニア、単純疱疹ウイルス、帯状水痘疱疹ウイルス (herpes zoster varicellous virus) およびライノウイルスによる感染を治療するのに有効であることが示されてきている。

### 【0153】

既定のターゲッティング配列が相同である標的部位は、IFN $\alpha$ 2遺伝子のコード領域内に、コード領域の上流にそしてすぐ隣接して、またはコード領域の上流にそして少し離れて存在していてもよい。例えば、2つのターゲッティング配列の第一のもの（または構築物中に1つのターゲッティング配列しかない場合、全ターゲッティング配列）は、IFN $\alpha$ 2コード配列の上流のゲノム領域由来であってもよい。例えば、このターゲッティング配列は、IFN $\alpha$ 2遺伝子のヌクレオチド-4074から-511に対応する、配列番号4の部分（例えば少なくとも20、50、100または1000の連続するヌクレオチド）を含んでもよい。2つのターゲッティング配列の第二のものは、コード配列自体の上流のゲノム領域を標識としてもよい。例えば、第二のターゲッティング配列は、その3'端に、IFN $\alpha$ 2コード配列の最初の数コドンに同一の外因性コード領域を含んでもよい。相同組換えに際して、外因性コード領域を、内因性IFN $\alpha$ 2コード配列の標的とされる部分と組み換える。IFN $\alpha$ 2を標的とするのに使用可能な配列は、その内容が完全に本明細書に援用される、米国特許出願第09/305,638号にさらに記載される。

### 【0154】

別の側面において、ターゲッティング配列は、ヒト顆粒球コロニー刺激因子 (GCSF) の部分に相同であってもよい。GCSFは、好中球/顆粒球細胞系譜に関係付けられる造血前駆細胞の増殖および分化を刺激するサイトカインである。GCSFは、化学療法が誘導する好中球減少の予防に、そして骨髓移植と関連



して、日常的に用いられる。慢性特発性および先天性好中球減少障害もまた、G C S F注射後、改善を示す。既定のターゲッティング配列が相同である標的部位は、G C S F遺伝子のエクソンおよび／またはイントロン内に、G C S Fコード領域の上流にそしてすぐ隣接して、またはG C S Fコード領域の上流にそして少し離れて存在していてもよい。

#### 【0155】

例えば、構築物中の2つのターゲッティング配列の第一のもの（または構築物中に1つのターゲッティング配列しかない場合、全ターゲッティング配列）は、G C S Fコード配列の上流のゲノム領域由来であってもよい。例えば、このターゲッティング配列は、ヒトG C S F遺伝子のヌクレオチド-6, 578から101に対応する、配列番号5の部分（例えば位-6, 578から-364（配列番号6）に対応する配列由来の少なくとも20、50、100、または1000の連続するヌクレオチド）を含んでもよい。構築物中の2つのターゲッティング配列の第二のものは、コード配列上流のゲノム領域を標的としてもよいし（例えば、やはり配列番号6の部分を含んでもよい）、あるいは遺伝子のエクソンまたはイントロンを標的としてもよい。G C S Fを標的とするのに使用可能な配列は、その内容が完全に本明細書に援用される、米国特許出願第09/305, 384号にさらに記載される。

#### 【0156】

##### 制御配列

DNA配列は制御配列を含むことが可能である。制御配列は、1以上のプロモーター（構成または誘導性プロモーター）、エンハンサー、UAS、足場付着領域またはマトリックス付着部位、陰性制御要素、転写因子結合部位、またはこれらの配列の組み合わせを含んでもよい。

#### 【0157】

制御配列は、産生される際、または個体に導入される際、細胞が、産物を発現するよう誘導可能であるように、誘導性プロモーターを含むことが可能であり、例えば細胞は該産物を発現しないが、発現するよう誘導することが可能である。制御配列は、制御配列の導入に際して、産物が発現されるように、誘導性プロ



モーターを含むことが可能である。制御配列は、細胞性またはウイルス配列であってもよい。こうした制御配列には、限定されるわけではないが、SV40初期または後期遺伝子、アデノウイルス主要後期遺伝子、マウスメタロチオネイン-Ⅰ遺伝子、伸長因子-1  $\alpha$  遺伝子、サイトメガロウイルス遺伝子、コラーゲン遺伝子、アクチン遺伝子、免疫グロブリン遺伝子、 $\gamma$  アクチン遺伝子、転写活性化因子YY1遺伝子、フィブロネクチン遺伝子、またはHMG-CoAレダクターゼ遺伝子の発現を制御するものが含まれる。制御配列は、TATAボックス、CCAATボックス、AP1、Sp1またはNF- $\kappa$ B結合部位などの転写因子結合部位を、さらに含むことが可能である。

### 【0158】

#### さらなるDNA配列要素

DNA配列はさらに、1以上のエクソンを含むことが可能である。エクソンは、RNAにコピーされ、そして成熟mRNA分子に存在するDNA配列である。エクソンは、1以上のアミノ酸をコードし、そして／またはアミノ酸を部分的にコードする（すなわちコドンの1または2塩基）DNAを含むことが可能である。あるいは、エクソンは、非コード領域、例えば5'非コード領域に対応するDNAを含む。外因性エクソンまたはエクソン類が1以上のアミノ酸および／またはアミノ酸の部分にコードする場合、DNA配列は、転写およびスプライシングに際して、読み枠が第二のエクソンまたは標的とされる遺伝子のコード領域とインフレームであるように、設計することが可能である。本明細書において、インフレームは、第一のエクソンおよび第二のエクソンのコード配列が、融合した際、第二のエクソンに由来するmRNAの部分の適切な読み枠を変化させない方式で、共にヌクレオチドに連結されることを意味する。

### 【0159】

標的とされる遺伝子の第一のエクソンが、翻訳を開始する配列ATGを含むならば、外因性エクソンは、好ましくはATGを含む。さらに、ATGを含む外因性エクソンは、第二のエクソンおよび標的とされる遺伝子の続くエクソンを含むmRNAの生じたコード領域がインフレームであるように、1以上のヌクレオチドをさらに含んでもよい。第一のエクソンがATGを含む、こうした標的とされ



る遺伝子の例には、ヒトエリスロポエチン、ヒト成長ホルモン、ヒトコロニー刺激因子ー顆粒球／マクロファージ（hGM-CSF）、およびヒトコロニー刺激因子ー顆粒球（hG-CSF）をコードする遺伝子が含まれる。

#### 【0160】

スプライスドナー部位は、1つのエクソンから別のエクソンへのスプライシングを指示する配列である。典型的には、第一のエクソンは、第二のエクソンの5'にあり、そして第一のエクソンの3'側で、第一のエクソンと重複し、そして隣接するスプライスドナー部位が、第二のエクソンの5'側にある第二のエクソンに隣接するスプライスーアクセプター部位を認識する。スプライスドナー部位は：第四および第五位にあるGUが必要とされる、(A/C)AG GU RAGU（Rはプリンヌクレオチドを示す）と示される特徴的なコンセンサス配列を有する可能性がある（Jackson (1991) Nucleic Acids Res. 19:3715-3798）。スプライスドナーコンセンサス部位の最初の3塩基は、エクソンの最後の3塩基である。スプライスドナー部位は、mRNAスプライシング経路内で、適切な反応を達成する能力によって、機能的に定義可能である。

#### 【0161】

非対スプライスドナー部位は、標的とされる配列に存在し、そしてDNA配列において、非対スプライスドナー部位の3'に配置されるスプライスーアクセプター部位が付随していない、スプライスドナー部位である。非対スプライスドナー部位は、内因性スプライスーアクセプター部位へのスプライシングを生じることが可能である。

#### 【0162】

スプライスーアクセプター部位は、スプライスドナー部位同様、1つのエクソンから別のエクソンへのスプライシングを指示する配列である。スプライシング装置は、スプライスドナー部位と組み合わせて作用し、スプライスーアクセプター部位を用いて、イントロンの除去を達成する。スプライスーアクセプター部位は：YYYYYYYYNYAG、ここでYはピリミジンいずれかを示し、そしてNはヌクレオチドいずれかを示す、として示される特徴的な配列を有す



る可能性がある (J a c k s o n (1991) N u c l e i c A c i d s R e s. 19:3715-3798)。

#### 【0163】

イントロンは、2つのエクソン間にあり、そしてmRNA分子の形成に際して、前駆体RNA分子から、スプライシングによって除去される、1以上のヌクレオチドの配列として定義される。

#### 【0164】

制御配列は、翻訳を開始するため、ATG開始コドンに連結することが可能である。所望により、CAP部位(制御領域と関連し、そして該領域に利用される、特定のmRNA開始部位)を制御配列およびATG開始コドンに連結することが可能である。あるいは、制御配列に関連し、そして該配列に利用されるCAP部位は、標的配列に含まれず、そして転写装置が新規CAP部位を提供する。CAP部位は、通常、TATAボックスのおよそ25ヌクレオチド3'に見出すことが可能である。スプライスドナー部位は、ATGにすぐ隣接して、例えば標的とされる遺伝子の第二のエクソンとインフレームであるために、外因性エクソンに、1以上のヌクレオチドの存在が必要とされない箇所に配置可能である。標的とされる遺伝子のコード配列とインフレームである、1以上のアミノ酸またはアミノ酸の部分にコードするDNAは、ATGの3'側にすぐ隣接して配置してもよい。こうしたものとして、スプライスドナー部位は、コードDNAの3'側にすぐ隣接して配置してもよい。

#### 【0165】

DNA配列のコード部分(例えばDNA配列のエクソン1)は、内因性タンパク質のものと同一の、1以上のアミノ酸、および/またはアミノ酸の部分にコードしてもよい。例えば、コードDNA配列は、目的の遺伝子の第一のエクソンに対応してもよい。あるいは、例えば、目的のタンパク質の第一のエクソンのアミノ酸が、該タンパク質の活性または活性類に重要でない場合、コードDNAは、目的のタンパク質の第一のエクソンと異なる、1以上のアミノ酸またはアミノ酸の部分にコードしてもよい。例えば、内因性ヒトエリスロポエチン(EPO)遺伝子への融合体を構築する場合、ヒト成長ホルモン(hGH)の第一のエクソン



をコードする配列が使用可能である。この例では、hGHエクソン1のEPOエクソン2への融合は、機能するハイブリッドシグナルペプチドの形成を生じる。しかし、コードされるアミノ酸がハイブリッドシグナルペプチドの機能を妨げない、ヒトまたは非ヒト起源のいかなるエクソンも使用可能である。

#### 【0166】

望ましい産物が、内因性タンパク質およびDNA配列中のコード配列の融合タンパク質である場合、細胞に取り込まれる外因性コードDNAは、内因性の標的とされる遺伝子の産物に融合しようとする、翻訳または転写産物に対応するcDNAの1以上のエクソンまたは配列をコードするDNAを含んでもよい。したがって、ターゲティングを用いて、2以上のタンパク質の構造的、酵素的あるいはリガンドまたは受容体結合特性を1つのポリペプチドに合併させる、キメラまたは多機能タンパク質を調製することが可能である。例えば、外因性DNA配列は、例えば標的とされるタンパク質の膜への係留、あるいは細胞分泌を提供するかまたは改善するシグナルペプチド、リーダー配列、酵素領域、膜貫通ドメイン領域、補助因子結合領域または他の機能領域をコードしてもよい。通常分泌されないが、シグナルタンパク質に融合させて、分泌を提供することが可能なタンパク質の例には、ドーパ・デカルボキシラーゼ、転写制御タンパク質およびチロシンヒドロキシラーゼが含まれる。

#### 【0167】

DNA配列は、天然に存在するか、あるいは遺伝子操作技術または合成法を用いて産生可能である供給源から得ることが可能である。

#### 標的配列

初代、二次または不死化細胞などの細胞にトランスフェクションした際、DNA配列は、望ましい産物、例えばタンパク質またはRNAの活性または機能部分の発現を調節することが可能である。DNA配列はまた、望ましい産物をコードしてもよい。産物は、例えば、ホルモン、サイトカイン、抗原、抗体、酵素、凝固因子、輸送タンパク質、受容体、制御タンパク質、構造タンパク質、転写因子、アンチセンスRNA、またはリボザイムであってもよい。さらに、産物は、天然に生じないタンパク質または核酸（すなわち融合タンパク質または核酸）であ



ってもよい。

### 【0168】

こうした産物には、エリスロポエチン、カルシトニン、成長ホルモン、インスリン、インスリノトロピン、インスリン様増殖因子、副甲状腺ホルモン、インターフェロン $\beta$ およびインターフェロン $\gamma$ 、神経増殖因子類、F S H  $\beta$ 、T G F  $\beta$ 、腫瘍壊死因子、グルカゴン、骨増殖因子-2、骨増殖因子-7、T S H  $\beta$ 、インターロイキン1、インターロイキン2、インターロイキン3、インターロイキン6、インターロイキン11、インターロイキン12、C S F-顆粒球、C S F-マクロファージ、C S F-顆粒球/マクロファージ、免疫グロブリン類、触媒性抗体類、プロテインキナーゼC、グルコセレブロシダーゼ、スーパーオキシドジスムターゼ、組織プラスミノゲン活性化因子、ウロキナーゼ、アンチトロンビンIII、DNアーゼ、 $\alpha$ -ガラクトシダーゼ、チロシンヒドロキシラーゼ、血液凝固因子V、血液凝固因子VII、血液凝固因子VIII、血液凝固因子IX、血液凝固因子X、血液凝固因子XI、アポリポタンパク質Eまたはアポリポタンパク質A-I、グロビン類、低密度リポタンパク質受容体、IL-2受容体、IL-2アンタゴニスト類、アルファ-1-アンチトリプシン、免疫反応修飾剤類、 $\beta$ -グルコセラミダーゼ、 $\alpha$ -イズロニダーゼ、 $\alpha$ -L-イズロニダーゼ、グルコサミン-N-スルファターゼ、 $\alpha$ -N-アセチルグルコサミニダーゼ、アセチル補酵素A： $\alpha$ -グルコサミド-N-アセチルトランスフェラーゼ、N-アセチルグルコサミン-6-スルファターゼ、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、N-アセチルガラクトサミン-6-スルファターゼ、および可溶性CD4が含まれる。

### 【0169】

#### 選択可能マーカーおよび増幅

ターゲッティング事象の同定は、1以上の選択可能マーカー遺伝子の使用によって、容易にすることが可能である。これらのマーカーは、DNA配列内に含まれてもよいし、または異なる構築物上に存在してもよい。選択可能マーカーは、2つのカテゴリー：陽性選択可能および陰性選択可能（言い換えると、陽性選択または陰性選択いずれかのためのマーカー）に分割可能である。陽性選択では、



陽性選択可能マーカーを発現する細胞は、選択剤（neo、キサンチン・グアニン・ホスホリボシルトランスフェラーゼ（gpt）、dhfr、アデニンデアミナーゼ（ada）、ピューロマイシン（pac）、ハイグロマイシン（hyg）、カルバミルリン酸シンターゼ、アスパラギン酸トランスカルバミラーゼ、およびジヒドロオロターゼグルタミンシンテターゼ（GS）をコードするCAD、多剤耐性1（mdr1）並びにヒスチジンD（hisD）など）での処理を生き抜くことが可能であり、ターゲッティング構築物が、宿主細胞ゲノムに組み込まれた細胞の選択を可能にする。陰性選択では、陰性選択可能マーカーを発現する細胞は、選択剤の存在下で破壊される。ターゲッティング事象の同定は、陰性選択可能マーカーが、外因性DNA配列に連結されるが、ターゲッティング配列に隣接するように、そして宿主細胞ゲノム中の配列との正しい相同組換え事象が陰性選択可能マーカーの安定な組込みを生じないように配置されるような、陰性選択の特性を示す1以上のマーカー遺伝子の使用によって、容易にすることが可能である（Mansourら（1988） Nature 336:348-352）。この目的に有用なマーカーには、単純疱疹ウイルスチミジンキナーゼ（TK）遺伝子、細菌gpt遺伝子、ジフテリア毒素、および細胞生存に必須な遺伝子をコードするmRNAのアンチセンスRNAまたはリボザイムが含まれる。

#### 【0170】

多様な選択可能マーカーを、初代、二次または不死化細胞に取り込んでもよい。例えば、薬剤耐性、栄養素要求性、細胞傷害剤に対する耐性または表面タンパク質の発現などの選択可能表現型を与える選択可能マーカーを用いてもよい。使用可能な選択可能マーカー遺伝子には、neo、gpt、dhfr、ada、pac、hyg、CAD、GS、mdr1およびhisDが含まれる。与えられた選択可能表現型は、レシピエント細胞を同定し、そして単離することを可能にする。

#### 【0171】

選択可能マーカーをコードする遺伝子（例えばada、GS、dhfrおよび多機能CAD遺伝子）は、増加したコピーの選択可能マーカーおよび隣接するゲノム配列を含む細胞の選択を可能にする、付加された特性を有する。この特徴は



、増加したコピーが望ましい、隣接または連結遺伝子のコピー数を有意に増加させるための機構を提供する。改善された選択特性を示すこれらの配列の突然変異型および増加したコピーを導く他の配列もまた、使用可能である。

### 【0172】

DNA配列中の構成要素の順序および数は、多様である可能性がある。例えば、順序は：第一のターゲッティング配列——選択可能マーカー——制御配列——エクソン——スプライスドナー部位——第二のターゲッティング配列であってもよいし、または代替物において、第一のターゲッティング配列——制御配列——エクソン——スプライスドナー部位——選択可能マーカーをコードするDNA——第二のターゲッティング配列であってもよい。構築物を安定に組み込む細胞は、選択剤での処理を生き抜くであろうし；安定にトランスフェクションされた細胞のサブセットは、相同組換え細胞であろう。相同組換え細胞は、PCR、サザンハイブリダイゼーションおよび表現型スクリーニングを含む、多様な技術によって、同定可能である。構築物の順序は：第一のターゲッティング配列——選択可能マーカー——制御配列——エクソン——スプライスドナー部位——イントロン——スプライスアクセプター部位——第二のターゲッティング配列であってもよい。

### 【0173】

あるいは、DNA配列中の構成要素の順序は、例えば：第一のターゲッティング配列——選択可能マーカー1——制御配列——エクソン——スプライスドナー部位——第二のターゲッティング配列——選択可能マーカー2であってもよいし、あるいは、第一のターゲッティング配列——制御配列——エクソン——スプライスドナー部位——選択可能マーカー1——第二のターゲッティング配列——選択可能マーカー2であってもよい。この配置では、選択可能マーカー2は、陰性選択の特性を示してもよい。すなわち、選択可能マーカー2の遺伝子産物は、選択可能マーカー2を発現している細胞を殺す剤（典型的には薬剤または代謝産物類似体）を含む適切な培地処方中の増殖によって、反対に選択可能である。選択可能マーカー1に隣接するターゲッティング配列と宿主細胞ゲノム中の相同配列間の組換えは、選択可能マーカー1の標的とされる組込みを生じ、一方、選



択可能マーカー2は組み込まれない。こうした組換え事象は、選択可能マーカー1で安定にトランスフェクションされているが、選択可能マーカー2で安定にトランスフェクションされていない細胞を生成し、そしてこうした細胞は、選択可能マーカー1を支持して選択する選択剤および選択可能マーカー2に反対して選択する選択剤を含む培地中の増殖によって、選択可能である。

#### 【0174】

DNA配列はまた、そのマーカーの増加したコピーを含む細胞の選択を可能にする、陽性選択可能マーカーも含むことが可能である。こうしたマーカーの増加したコピーは、隣接するDNA配列の同時増幅を生じる。例えば、構成要素の順序は：第一のターゲティング配列——コピー数を増加させる陽性選択可能マーカー——第二の選択可能マーカー（所望による）——制御配列——エクソン——スプライスドナー部位——第二のターゲティングDNA配列であってもよい。活性化される遺伝子は、選択可能マーカー遺伝子の増加したコピーを含む細胞が、適切な選択可能剤の存在下で、細胞を培養することによって、選択可能である特性を有する、選択可能マーカー遺伝子の包含によって、さらに増加させることが可能である。活性化される内因性遺伝子は、選択可能マーカー遺伝子とタンデムに増加されるであろう。活性化される内因性遺伝子を多コピー含む細胞は、非常に高レベルの望ましいタンパク質を産生することが可能であり、そして *in vitro* タンパク質産生および遺伝子治療に有用である。

#### 【0175】

選択可能および他のマーカー遺伝子は、互いにすぐ隣接している必要はない。

#### DNA配列／相同組換え増進剤／非相同端連結阻害剤複合体

二本鎖DNA配列および選択される標的DNA、例えば細胞中の染色体DNA間の相同組換えは、相同組換えを増進する剤、例えばRad52タンパク質、および非相同端連結を阻害する剤、例えばKu不活性化剤（例えば抗Ku抗体）を、DNA配列および標的とされる部位に十分にごく近接して提供することによって、促進可能である。「十分にごく近接」は、本明細書において、相同組換え増進剤および／または非相同端連結を阻害する剤の濃度が、標的とされる部位の改変、例えばDNA配列および標的配列間の相同組換えを、より高い率で提供する



のに十分である、相同組換え増進剤または非相同端連結を阻害する剤、あるいは両方の導入を指す。いくつかの方法を用いて、DNA配列、相同組換え増進剤、および非相同端連結を阻害する剤の、互いにごく近接した導入を提供することが可能である。これらの化合物を互いに、そして標的DNAに、ごく近接して投与することによって、Rad52およびKu不活性化分子、例えば抗Ku抗体などの化合物の活性を、相同組換え部位に局在させる。例えば、Ku活性の局所阻害は、Ku活性の全細胞阻害より好ましい可能性がある。

#### 【0176】

DNA配列、相同組換え増進剤、および非相同端連結を阻害する剤の近接は、これらの要素を複合体の一部として導入することによって、維持可能である。例えば、DNA-タンパク質複合体が使用可能である。DNA-タンパク質複合体の中心は、標的DNAの選択される部位に導入しようとする二本鎖DNA配列で構成されていてもよい。相同組換え増進剤、例えばRad52タンパク質またはその断片は、DNA配列上、例えばDNA配列の全配列または端のみに、例えばDNA配列の一本鎖突出端の少なくとも部分上に、付着して、例えばコーティングされていてもよい。DNA-タンパク質複合体は、DNA配列または相同組換え増進剤いずれかに共有結合している、非相同端連結を阻害する剤、例えば抗Ku抗体などのKu不活性化剤をさらに含むことが可能である。非相同端連結を阻害する剤はまた、DNA配列または相同組換え増進剤に非共有結合していてもよい。

#### 【0177】

化合物はまた、DNA配列、相同組換え増進剤および非相同端連結を阻害する剤を、リポソームまたは小胞中に提供することによって、互いにごく近接して維持することも可能である。例えば、リポソーム懸濁物はまた、これらの要素の薬学的に許容しうるキャリアーとしても、使用可能である。リポソーム懸濁物は、例えば、米国特許第4,522,811号に記載されるように、当業者に知られる方法にしたがって、調製可能である。

#### 【0178】

DNA配列、相同組換え増進剤および非相同端連結を阻害する剤はまた、細胞



にマイクロインジェクションすることが可能な混合溶液の一部であってもよいし、またはこれらの化合物3つすべてが、細胞に同時に存在するように、これらの化合物各々を、他のものに続いて、迅速に導入してもよい。1以上のこれらの化合物を導入する他の方法には、受容体仲介搬送、エレクトロポレーションおよびリン酸カルシウム沈殿が含まれる。

#### 【0179】

##### 細胞

トランスフェクションしようとする初代および二次細胞は、多様な組織から得ることが可能であり、そして培養中で維持および増殖可能な細胞種を含む。例えば、トランスフェクション可能な初代および二次細胞には、線維芽細胞、ケラチン形成細胞、上皮細胞（例えば乳腺上皮細胞、腸上皮細胞）、内皮細胞、グリア細胞、神経細胞、血液の形成された要素（例えばリンパ球、骨髄細胞）、筋細胞およびこれらの体細胞種の前駆体が含まれる。初代細胞は、好ましくは、トランスフェクション初代または二次細胞が投与される個体から得られる（すなわち自己細胞）。しかし、初代細胞は、同一種（すなわち同種異系細胞）または別の種（すなわち異種細胞）（例えばマウス、ラット、ウサギ、ネコ、イヌ、ブタ、ウシ、小鳥、ヒツジ、ヤギ、ウマ）のドナー（レシピエント以外）から得てもよい。

#### 【0180】

脊椎動物、好ましくは哺乳動物起源の初代または二次細胞は、外因性DNA配列、例えば療法タンパク質をコードする外因性DNA配列でトランスフェクションされ、そして*in vitro*および*in vivo*両方で、延長された期間、コードされる療法タンパク質を、安定に、そして再現可能に産生することが可能である。さらに、トランスフェクション初代および二次細胞は、生理学的に適切なレベルで、*in vivo*で、コードされる産物を発現可能であり、移植後に細胞を回収し、そして再培養に際して、増殖させ、そしてその移植前特性を示すことが可能である。

#### 【0181】

あるいは、脊椎動物、好ましくは哺乳動物起源の初代または二次細胞は、制御



配列を含む外因性DNA配列でトランスフェクション可能である。こうした制御配列の例には、1以上の：プロモーター、エンハンサー、UAS、足場付着領域または転写結合部位が含まれる。ターゲッティング事象は、DNA配列の制御配列の挿入を生じ、標的とされる内因性遺伝子をその調節下に置くことが可能である（例えば、プロモーターまたはエンハンサー、あるいは両方を、内因性遺伝子または制御領域の上流に挿入することによって）。所望により、ターゲッティング事象は、遺伝子の組織特異的陰性制御配列の欠失などの、内因性制御配列の欠失を、同時に生じることが可能である。ターゲッティング事象は、存在する制御配列を置換することが可能であり；例えば組織特異的エンハンサーを、天然存在要素より、より広いかまたはそれと異なる細胞種特異性を有するエンハンサー、あるいは対応する非トランスフェクション細胞と異なる制御または誘導パターンを示すエンハンサーで置換することが可能である。これに関連し、天然存在配列が欠失され、そして新規配列が付加される。あるいは、内因性制御配列は、除去または置換されないが、内因性制御要素内に外因性配列をターゲッティングすることによるなど、ターゲッティング事象によって、破壊されるかまたは無能にされる。相同組換えによる制御配列の導入は、通常発現しない療法タンパク質を発現する初代または二次細胞を生じる可能性がある。さらに、制御配列のターゲッティングされる導入は、療法タンパク質を作成するかまたは含むが、正常より少量（生理学的に正常な、より低いレベル未満の量）であるか、または欠損型である細胞に、そして生理学的に正常なレベルで療法タンパク質を作成するが、その含量または産生を増大させるかまたは増進しようとする細胞に、使用可能である。

#### 【0182】

トランスフェクション初代または二次細胞はまた、選択可能表現型を与え、その同定および単離を容易にする、選択可能マーカーをコードするDNA配列もまた、含んでもよい。安定にDNA配列を発現する、トランスフェクション初代、二次細胞を産生する方法、こうしたトランスフェクション細胞のクローン性細胞株および不均質細胞株、クローン性および不均質細胞株を産生する方法、並びにトランスフェクション初代または二次細胞の集団の使用を通じて、異常なまたは



望ましくない状態を治療するかまたは予防する方法が、本発明の一部である。

### 【0183】

初代または二次細胞のトランスフェクション、相同組換えおよびクローン性または不均質細胞株の産生

脊椎動物組織は、パンチ生検または目的の初代細胞種の組織供給源を得る、他の手術法などの標準的方法によって、得ることが可能である。例えば、パンチ生検は、線維芽細胞またはケラチン形成細胞の供給源として皮膚を得るのに用いる。初代細胞の混合物は、酵素消化または外植などの既知の方法を用いて、組織から得る。酵素消化を用いた場合、コラゲナーゼ、ヒアルロニダーゼ、ディスパーゼ、プロナーゼ、トリプシン、エラスターゼおよびキモトリプシンなどの酵素が使用可能である。

### 【0184】

生じた初代細胞混合物を直接トランスフェクションしてもよいし、またはトランスフェクションを行う前に、まず培養し、培養プレートから除去し、そして再懸濁してもよい。初代細胞または二次細胞は、所望により選択可能マーカーをコードするDNAを含む、ゲノムに導入しようとするDNAと合わせ、そしてトランスフェクションを達成するために処理する。さらに、単独でまたは複合体の一部として、Rad52タンパク質またはその断片、およびKu不活性化分子、例えば抗Ku抗体と、初代または二次細胞を合わせる。

### 【0185】

トランスフェクション初代または二次細胞は、エレクトロポレーションによって、作成可能である。エレクトロポレーションは、適切な電位および電気容量（および対応する時定数）で行い、初代または二次細胞へのDNA構築物（類）の進入を生じる。エレクトロポレーションは、広い範囲の電位（例えば50から2000ボルト）および対応する電気容量に渡って実行可能である。通常、およそ0.1から500  $\mu$ gの総DNAを用いる。

### 【0186】

好ましくは、初代または二次細胞は、マイクロインジェクションを用いてトランスフェクションする。あるいは、リン酸カルシウム沈殿、修飾リン酸カルシウ



ム沈殿およびポリブレン沈殿、リポソーム融合および受容体仲介遺伝子搬送などの既知の方法を用いて、細胞をトランスフェクションすることが可能である。安定なトランスフェクション細胞を単離し、そして培養条件下で、そして十分な時間、培養し、そして継代培養して、安定なトランスフェクション二次細胞を増殖させ、そしてトランスフェクション二次細胞のクローン性細胞株を産生する。あるいは、1より多いトランスフェクション細胞を培養し、そして継代培養し、不均質な細胞株の産生を生じる。

#### 【0187】

トランスフェクション後、当該技術分野に知られるように (C a p e c c h i (1989) S c i e n c e 244:1288-1292)、相同組換えを可能にする条件下で、細胞を維持する。

#### 【0188】

相同組換え初代または二次細胞は、十分な数の倍加を経て、有効量で、個体に療法タンパク質を提供するのに十分なサイズのクローン性細胞株または不均質細胞株いずれかを生じることが可能である。一般的に、例えば、 $0.1\text{ cm}^2$ の皮膚を生検で得て、そしてこれは100,000細胞を含むと考えられ；1つの細胞を用いて、クローン性細胞株を産生し、そしておよそ27の倍加を経て、1億の相同組換え二次細胞を産生する。不均質細胞株を、およそ100,000細胞の元来の相同組換え集団から産生しようとする場合、1億の細胞を産生するのに10の倍加しか必要とされない。

#### 【0189】

相同組換えクローン性または不均質細胞株に必要とされる細胞の数は多様であり、そして限定されるわけではないが、相同組換え細胞の使用、細胞における外因性DNA配列の機能レベル、細胞における改変されたDNA配列の機能レベル、相同組換え細胞の移植部位（例えば使用可能な細胞数は、移植の解剖学的部位によって制限される）、並びに患者の年齢、表面面積、および臨床的状态を含む、多様な要因に応じる。これらの要因を視野に置くと、成長ホルモン単独欠損症を有する、そうでなければ健康な10kgの患者にヒト成長ホルモンの療法レベルを搬送するため、およそ1から5億の相同組換え線維芽細胞が必要であろう（



これらの細胞の体積は、ほぼ、患者の親指のごく先端程度である）。

#### 【0190】

いくつかの方法を用いて、本明細書に記載される方法が、細胞において相同組換えを増進する有効性を決定することが可能である。例えば、細胞、例えばヒト細胞における非保存的置換を検出するため、実験系を設計することが可能である。置換は、XhoI部位の一部である、HPRT遺伝子のエクソン3のCGAコドンでのCからTの置換であってもよい。この突然変異は、TGA終結シグナルを生成し、これは、6-チオグアニン（6-TG）に耐性とスコア付けされる、HPRT陰性表現型を生じる。この突然変異はまた、対応するXhoI部位の欠失を伴う。簡潔には、CからTの置換を含むDNA配列は、相同組換えを増進する剤およびKuを不活性化する剤を含む複合体の一部として、ヒト線維芽細胞にマイクロインジェクションすることによって、導入可能である。細胞を培養し、そして6-TGを含む培地上に細胞を導入する前に、増殖させる。その後、6-TG耐性クローンをスコア付けし、突然変異DNA配列の存在を決定する。相同組換え事象の存在は、HPRT特異的プローブを用いて、XhoI消化ゲノムDNAをサザンブロット解析することによって、検出可能である。結果はまた、突然変異DNA配列が、相同組換えを増進する剤およびKuを不活性化する剤の非存在下に導入される、対照細胞と比較することも可能である。

#### 【0191】

##### 相同組換え二次細胞のクローン性細胞株または不均質細胞株の移植

上述のように産生された相同組換え細胞は、既知の方法を用いて、療法タンパク質を搬送しようとする個体に導入可能である。その後、クローン性細胞株または不均質細胞株を、既知の方法を用いて、多様な投与経路および多様な部位（例えば腎被膜下、皮下、中枢神経系（クモ膜下腔内を含む）、血管内、肝臓内、内臓内、腹腔内（大網内を含む）、または筋肉移植）を用い、個体に導入する。個体に移植されたら、相同組換え細胞は、外因性合成DNAにコードされる療法産物を産生するか、または相同組換え細胞は、外因性制御配列調節下で、内因性DNA配列にコードされる療法タンパク質を発現する。例えば、血液に通常見られる因子IXの欠損によって引き起こされる出血障害である、血友病Bと診断され



ている個体は、遺伝子治療の候補である。患者に小皮膚生検を行う；これは、外来に基づいて行うことが可能な、単純な処置である。ほぼマッチの頭の大きさの皮膚片を、例えば腕の下から取り、そしてこれは除去するのに約1分を要する。試料をプロセッシングし、患者細胞（この場合、線維芽細胞）を単離し、そして失われた因子 I X を産生するように、遺伝子操作する。患者の年齢、体重、および臨床的状态に基づいて、必要とされる数の細胞を、大規模培養で増殖させる。全過程は、4－6週間を要するはずであり、そしてこの期間の終わりに、再び外来で（例えば患者の皮膚下に、これらを注入して戻すことによって）、遺伝子操作された適切な数の細胞を、個体に導入する。患者はここで、自身の因子 I X を産生することが可能であり、そしてもはや血友病患者ではない。

#### 【0192】

同様のアプローチを用いて、他の異常または疾患を治療することが可能である。例えば、低身長は、ヒト成長ホルモンを発現する初代または二次細胞を移植することにより、個体にヒト成長ホルモンを投与することによって、治療可能である。

#### 【0193】

この例が示唆するように、使用する細胞は、一般的に、患者特異的に遺伝子操作された細胞であろう。しかし、同一種の別の個体または異なる種から細胞を得ることが可能である。こうした細胞の使用は、免疫抑制剤の投与、組織適合性抗原の改変、または移植細胞の拒絶を防ぐ障壁装置の使用を要する可能性がある。

#### 【0194】

多くの疾患では、これは、単回治療であろうし、そして他のものでは、多数の遺伝子治療措置が必要とされるであろう。

トランスフェクション初代または二次細胞は、単独で、またはレシピエント被験者における細胞に対する免疫反応を阻害するための障壁または剤と組み合わせて、投与可能である。例えば、免疫抑制剤を被験者に投与し、被験者における正常反応を阻害するか、または該反応と干渉することが可能である。好ましくは、免疫抑制剤は、被験者において、T細胞／またはB細胞活性を阻害する免疫抑制薬剤である。こうした免疫抑制薬剤の例が商業的に入手可能である（例えばシク



ロスボリンAは、S a n d o z C o r p.、ニュージャージー州イーストハノーバーから商業的に入手可能である）。

#### 【0195】

免疫抑制剤、例えば薬剤は、望ましい療法効果（例えば細胞の拒絶の阻害）を達成するのに十分な投薬量で、被験者に投与可能である。免疫抑制剤の投薬範囲は、当該技術分野に知られる。例えばF r e e dら（1992） N. E n g l. J. M e d. 327:1549; S p e n c e rら（1992） N. E n g l. J. M e d. 327:1541; W i d n e rら（1992） N. E n g l. J. M e d. 327:1556を参照されたい。投薬値は、個体の疾患状態、年齢、性別、および体重などの要因にしたがって、多様である可能性がある。

#### 【0196】

被験者においてT細胞活性を阻害するのに使用可能な別の剤は、抗体、あるいはその断片または誘導体である。i n v i v oでT細胞を枯渇させるかまたは隔絶することが可能な抗体が当該技術分野に知られる。ポリクローナル抗血清、例えば抗リンパ球血清が使用可能である。あるいは、1以上のモノクローナル抗体が、使用可能である。好ましいT細胞枯渇抗体には、細胞表面上のC D 2、C D 3、C D 4、C D 8、C D 40、C D 40リガンドに結合するモノクローナル抗体が含まれる。こうした抗体が当該技術分野に知られ、そして例えばアメリカン・タイプ・カルチャー・コレクションから商業的に入手可能である。ヒトT細胞上のC D 3に結合する好ましい抗体は、O K T 3（A T C C C R L 8001）である。

#### 【0197】

レシピエント被験者内でT細胞を枯渇させるか、隔絶するかまたは阻害する抗体は、移植に際して細胞の拒絶を阻害する用量で、適切な期間、投与することが可能である。抗体は、好ましくは、薬学的に許容しうるキャリアーまたは希釈剤（例えば生理食塩水溶液）中で、静脈内投与する。

#### 【0198】

レシピエント被験者において、細胞に対する免疫反応と干渉するか、または該



反応を阻害する別の方法は、免疫障壁を使用することである。「免疫障壁」は、本明細書において、被験者における投与された細胞および免疫反応に関与する細胞間の障壁として役立つ装置を指す。例えば、細胞は、移植可能装置中で、投与することが可能である。移植可能装置は、半透性障壁、すなわち栄養素および産物が障壁内および障壁外に拡散するが、より大きい免疫系構成要素、例えば抗体または補体の進入を妨げる障壁の内部に含まれる細胞を含んでもよい。移植可能装置は、典型的には、マトリックス、例えばヒドロゲル、または細胞を配置するコアを含む。所望により、半透性コーティングは、ゲルを封入してもよい。ゲルコア内に配置される場合、投与される細胞は、免疫系の細胞から隔離されているはずであり、そして宿主の細胞および細胞傷害性抗体から覆い隠されているはずである。好ましくは、PLLまたはPLOなどの選択透過性コーティングを用いる。コーティングはしばしば、レシピエントの免疫系構成要素が進入して、そして移植可能装置内の細胞を破壊するのを妨げる、小孔を有する。

#### 【0199】

細胞を被包する多くの方法が当該技術分野に知られる。例えば、産生用の細胞を被包するため、水可溶性ゴム(gum)を用いて、半透性水不溶性ゲルを得る被包および他の被包法が、米国特許第4,352,883号に開示される。使用可能な他の移植可能装置は、米国特許第5,084,350号、米国特許第5,427,935号、1995年7月27に公表されたWO 95/19743、米国特許第5,545,423号、米国特許第4,409,331号、米国特許第4,663,286号、および欧州特許第301,777号に開示される。

#### 【0200】

##### 相同組換え初代および二次細胞および細胞株の使用

相同組換え初代および二次細胞または細胞株は、酵素、ホルモン、サイトカイン、抗原、抗体、凝固因子、アンチセンスRNA、制御タンパク質、転写タンパク質、受容体、構造タンパク質、新規(非最適化)タンパク質および核酸産物、並びに操作DNAなどの、療法タンパク質のためのビヒクルまたは搬送系として、広い適用可能性を有する。例えば、相同組換え初代または二次細胞を用いて、限定されるわけではないが、エリスロポエチン、カルシトニン、成長ホルモン、



インスリン、インスリノトロピン、インスリン様増殖因子、副甲状腺ホルモン、 $\alpha$ 2-インターフェロン (IFNA2)、 $\beta$ -インターフェロン、 $\gamma$ -インターフェロン、神経増殖因子類、FSH $\beta$ 、TGF- $\beta$ 、腫瘍壊死因子、グルカゴン、骨増殖因子-2、骨増殖因子-7、TSH- $\beta$ 、インターロイキン1、インターロイキン2、インターロイキン3、インターロイキン6、インターロイキン11、インターロイキン12、CSF-顆粒球 (GCSF)、CSF-マクロファージ、CSF-顆粒球/マクロファージ、免疫グロブリン類、触媒性抗体類、プロテインキナーゼC、グルコセレブロシダーゼ、スーパーオキシドジスムターゼ、組織プラスミノゲン活性化因子、ウロキナーゼ、アンチトロンビンIII、DNアーゼ、 $\alpha$ -ガラクトシダーゼ、チロシンヒドロキシラーゼ、血液凝固因子V、血液凝固因子VII、血液凝固因子VIII、血液凝固因子IX、血液凝固因子X、血液凝固因子XI、アポリポタンパク質E、アポリポタンパク質A-I、グロビン類、低密度リポタンパク質受容体、IL-2受容体、IL-2アンタゴニスト類、 $\alpha$ -1-アンチトリプシン、免疫反応修飾剤類、 $\beta$ -グルコセラミダーゼ、 $\alpha$ -イズロニダーゼ、 $\alpha$ -L-イズロニダーゼ、グルコサミン-N-スルファターゼ、 $\alpha$ -N-アセチルグルコサミニダーゼ、アセチル補酵素A： $\alpha$ -グルコサミン-N-アセチルトランスフェラーゼ、N-アセチルグルコサミン-6-スルファターゼ、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、N-アセチルガラクトサミン-6-スルファターゼ、および可溶性CD4を含む療法タンパク質を供給することが可能である。

### 【0201】

本明細書に引用されるすべての特許および参考文献は、その全体が本明細書に援用される。他の態様が請求項の範囲内である。

### 【配列表】



## SEQUENCE LISTING

&lt;110&gt; Evguenii Ivanov

&lt;120&gt; METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION

&lt;130&gt; 10278/016001

&lt;160&gt; 9

&lt;170&gt; FastSEQ for Windows Version 3.0

&lt;210&gt; 1

&lt;211&gt; 7622

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 1

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ggatccgaga  | acatagaagg  | agcaggtaat  | ttatcaaggc  | atgaacacgg  | gtgcttaatt  | 60   |
| tcttatattg  | aggccaggca  | tggtggctca  | cacctgtaat  | cccaacactt  | taggaagcoa  | 120  |
| aggtgggtgg  | attgottgag  | tctaggattt  | tgagaccagc  | ctggccaaca  | tggcgaaato  | 180  |
| ctgtctctac  | taaaaatact  | aaaattaaoc  | agtcattggt  | gtgggtgtgc  | tttagtpoca  | 240  |
| gctactctgg  | tggttgaggc  | acaagaatca  | cttgaacctg  | ggaggcagag  | gttgcagtga  | 300  |
| gctgagactg  | tgccaottca  | ctccagcctg  | ggtgacagag  | taagattctg  | totoaaaaaa  | 360  |
| tatgtatata  | tacacaceta  | taatagatac  | ataaaotat   | atcacatata  | aatatataaa  | 420  |
| tatatataat  | atatataata  | tataaacata  | tataaatata  | tatatatata  | tatatatata  | 480  |
| tatataaaoc  | aaacataaag  | gaataatatt  | gggggaaaa   | cttcataaat  | gaaagaacaa  | 540  |
| cataggotgt  | tgagtatatg  | cacagaaatt  | caagagatct  | tccagcaatt  | gaagacattg  | 600  |
| gtttaccaga  | attcacaana  | gaagtoagct  | gtgcatttaa  | agtagaatgt  | gatgagtgtt  | 660  |
| acaaactgag  | taggaactgg  | gaactaagga  | agcgttaagc  | agaaagtgtc  | gaactgagag  | 720  |
| ttggggcatt  | gaggctgtgt  | aaggcagggt  | aagtgaatgt  | ctcctagaag  | ctaaccttaa  | 780  |
| atggagtttt  | gaagtacttg  | taggagtagc  | ttaggtgaaa  | agaagaggag  | aaacatgtat  | 840  |
| caggcagagg  | gaotagaacc  | ttattacott  | caaagaagaa  | gcaaaaagaa  | tacatgtgac  | 900  |
| tttgagggtg  | tgagggtgac  | tttaagccaa  | tataggtgaa  | tttgacatag  | gaattcccta  | 960  |
| aataatgtto  | ggtcatttgt  | taaatattga  | gtgatataac  | actgtattaa  | agcccaagag  | 1020 |
| ttgcttttat  | atagaaaagaa | gaaaaaagcc  | caagagagtt  | ttatttctag  | agggaaatatt | 1080 |
| ttctagaaat  | aaaggaaggt  | gtatocagcca | gtttctagtc  | agggaaaacag | aaatoacacc  | 1140 |
| tgatatgcaa  | aataagaggaa | aatoagggaa  | ttcatttaato | cagagatttg  | gttgctcaag  | 1200 |
| tattagattg  | ctgaaaagoc  | agacagggaa  | tatgaggvaa  | tcagagataa  | gtattagtga  | 1260 |
| caagctccat  | ttatgtgcag  | gattggaggg  | acataggtgg  | ggttcccaga  | agocagaagg  | 1320 |
| tgagaccacc  | tagcagaagc  | tcaaacccaca | gctgggggtt  | octcacaaaa  | gctgggacca  | 1380 |
| ccaggaggag  | ctgtccaatg  | ggatctggag  | ccagggagat  | catgcagtca  | ctaacaggaa  | 1440 |
| gggaagcaga  | atgtaaaagg  | tagagagaaa  | tactccaact  | gcttccttgc  | attcaacttc  | 1500 |
| caatctocat  | tcacaaaagg  | aaaaacctgc  | taatacagca  | gagtgggaaa  | agcagcctgc  | 1560 |
| oaaggctcct  | tctcccacaa  | aacagagoc   | aaaaccagc   | aaaaacaagg  | aatgoatttg  | 1620 |
| atagcaaaac  | ggctatggac  | caacccaaca  | taaaagaaat  | gatgagtgat  | ttcttttttc  | 1680 |
| atttggttca  | agaaaagtat  | ttcagtaact  | attatgtaac  | agaaattota  | tttatttttg  | 1740 |
| ggaattcaaa  | ggtgaataaa  | aaagaactct  | aaatttttat  | caataaaata  | tttcaaaaac  | 1800 |
| ctcaatgaga  | gtaattggcat | taactagcaa  | atatgtcaat  | gagatgagct  | agccataaga  | 1860 |
| ggottagaat  | tgagagaaag  | gtctgggggc  | ctcttgacag  | gccaaattca  | gagctgtttg  | 1920 |
| tgggaatctc  | tgacctaaat  | gcagggtgaa  | atataaatat  | gggcatttag  | aatagtggcc  | 1980 |
| caaaactttg  | atgattttot  | tcttggggtc  | totocaatta  | atgggattga  | tgagaactgt  | 2040 |
| agaccactga  | ggtcaccatg  | gctcaatgaa  | tagtccoctg  | gctttggagt  | caaaactgac  | 2100 |
| tgaatatgaa  | ccccagottt  | gctacttaoa  | ggttgcattt  | atcctcagtt  | ttctcatctt  | 2160 |
| tcaaaagaaga | acagtaactt  | ctttaaaagg  | ttattgtagg  | ctgggtgcag  | tggtccacgc  | 2220 |
| ctgtaatcgc  | agcactttgg  | gaggcggagg  | ctagtggatc  | acttgaggcc  | aggagttaga  | 2280 |
| aactagcctg  | gccaacatgg  | tgaaactctg  | tctctacaaa  | aagaaattta  | aaaaattttg  | 2340 |
| ctgggtgtgg  | tggaacacac  | ctggaattcc  | agctacctgg  | gaggccgagg  | catgagcatc  | 2400 |
| acttgagtct  | ggaaagcaga  | gggttgcaat  | gagccaagat  | tgtaccactg  | tactcaagcc  | 2460 |
| tgggtgacac  | agtgagacot  | tgtctaaaaa  | aaaaaaagg   | tattgtgtta  | ttgtaaatat  | 2520 |
| tgtatatgaa  | ottctattta  | acatgtttag  | ttaaatgcct  | gtgtaattgt  | caaatgtgtg  | 2580 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| cttctagctc  | actgcacaga  | caaaactgat  | tcactgaaat  | catggaattg  | cagcaaaaga  | 2640 |
| caaatctaat  | taatgtagg   | caaaacggag  | gactggagtt  | attattcaaa  | tcagctccc   | 2700 |
| tgaanaactca | gaggctagg   | ttttatggat  | aatttgggtg  | gcaggggact  | aggggaatggg | 2760 |
| tgctgctgat  | tggttgggga  | atgaaatagt  | aagattgtgg  | aaaactgtcc  | tccttcattg  | 2820 |
| agtcagcttc  | oggggtgtagg | ccacacgacc  | agttgagtc   | tgaagcatgc  | gtccaaagtgg | 2880 |
| agtcagtttg  | ttgocagaat  | gcaaaagcct  | gaaaaatgtc  | tcaaatgatc  | aaactgtaggc | 2940 |
| tccacaataa  | tgtattattc  | tataggagca  | attggggaag  | taacaaatct  | tgtgacctct  | 3000 |
| ggacacataa  | ctcctgaact  | agtaagggat  | tataaaaaac  | atgocataat  | ottatcagaa  | 3060 |
| ttcaggtccc  | cccataatcc  | taatctocac  | gcattttcatt | tgtttagaaa  | ggccattttc  | 3120 |
| agtccttgag  | caaggagggg  | gttagtttta  | ggataggact  | attatccttg  | cttctgttaa  | 3180 |
| ctataaacta  | aattcctccc  | atgggttagct | tggtctacac  | ctaaagatga  | gtgagaaacag | 3240 |
| ccagcctgtg  | aggctagagg  | caagatggag  | tcagccatgc  | tagattttatc | tcactgtcat  | 3300 |
| aacctttgca  | aaggcagttt  | ccactgggac  | ataggaggta  | ctcaatgaaa  | aagaagctat  | 3360 |
| taataattaaa | atttttaaaa  | tgaatttaag  | gaaactaatc  | tatgtacata  | ttagtcat    | 3420 |
| aaacaaagtg  | gttcattttac | attcacacaa  | ataaatcttg  | tgattatata  | taggtaatat  | 3480 |
| gaaaaacttt  | gttttctttc  | ataatacaag  | gtattagcaa  | tagatatagt  | aatgttagca  | 3540 |
| ttcctttgga  | aaaaatgaaa  | agattttata  | ttttccaaga  | atcatttagta | tttttattta  | 3600 |
| atatacatata | tataaaattt  | attcattcta  | taacttggaa  | atatgcttgc  | ttaccatata  | 3660 |
| ctgacagatt  | tcacaaatatt | ctataactca  | caatattcat  | ttacataaat  | attgatttgg  | 3720 |
| taottacaat  | gtgtactgct  | atgctaagtt  | ttgtctttgt  | caaacatatt  | ttataaaatc  | 3780 |
| ataatcctag  | atgaatccaa  | otttttggtaa | cccagctgoc  | tgaacccctg  | ctgttaacag  | 3840 |
| gcaaatgtgtg | gtaggtagag  | atctataact  | accactttcc  | tctacccacc  | agcatctgca  | 3900 |
| cccaccaccc  | ctccccaccc  | accatttatct | ataccaaoca  | cccctcccaa  | cttacacagca | 3960 |
| totgacacca  | ccacaccgco  | ccccaccac   | catgtacact  | cactacacct  | tcagccatc   | 4020 |
| accactctgca | cccatcactc  | ctcccatcc   | acaagcatct  | gcacccacca  | catttcccta  | 4080 |
| cctaccagca  | ctttcactca  | ccactctctc  | accaccagc   | atctgcaccc  | acaaccctc   | 4140 |
| ctcaaccacc  | agagtctgca  | tcctacacac  | ttgcccactc  | gctagcatct  | gcaccatcaa  | 4200 |
| gctctgctct  | cttgctaat   | acgggatgag  | ctctccatgg  | ttctgcttaa  | agacaatgct  | 4260 |
| ttcaatctct  | ttctataacc  | catttctctt  | tacotcttca  | agtaacactc  | agaactctct  | 4320 |
| tctctctctg  | ataccaactt  | tttccacttt  | actcaatcat  | tcctatcacc  | atacaaaagt  | 4380 |
| gtttattttc  | cccatcttaa  | agttaaaaat  | caaaagaaaa  | ttgtctgctg  | ccaggcagcg  | 4440 |
| ttgtctoaogc | ctgtaatccc  | aacactttgg  | gaggccaaagg | aggggtggat  | gacttaaggt  | 4500 |
| taggagttca  | agaccagcct  | ggccaacatg  | gtgaaaccca  | tcctactata  | aaatacaaaa  | 4560 |
| attagccagg  | catggttgca  | catgcctgta  | gtctcaggtc  | cttgaggaggc | tgaggccaga  | 4620 |
| gaatggcttg  | aaocccggag  | gcagaggttg  | cagtgagcog  | agattgtgct  | cttgcaotcc  | 4680 |
| agcctgggtg  | acagagtgag  | actccatctc  | aaaaataaaa  | aataaaaata  | aaacaaaaga  | 4740 |
| aagttattttt | tacccaacat  | ccacattaac  | caaataccac  | tttctttatt  | gatotttgta  | 4800 |
| aaaaaaaagt  | cttgaaaaaa  | ttgtctatat  | tcactatgac  | ttatctctct  | caaatcactt  | 4860 |
| aaacacatac  | caatcaggtt  | ttgttttca   | gtactccaaa  | gtaactttta  | cagccaagga  | 4920 |
| cagtagecga  | ctttacatcg  | catatgcatt  | gtgaagttct  | tgatctctat  | cttacttaac  | 4980 |
| ctgtcagcag  | tatctgacac  | aggtgtcact  | ggctctctcc  | tgagatgctc  | tccttatttg  | 5040 |
| gctttggggga | caocatattc  | tcctctctcc  | tactttctct  | aatggccctc  | ctcagctctc  | 5100 |
| tttggaagaa  | ggaaaaagaa  | acttcattat  | ctcctggatg  | tagtacaac   | aaactcaagct | 5160 |
| caacatgtgc  | atactgaact  | coatttctct  | ttcccaaaat  | tcgacattta  | cagccatccc  | 5220 |
| ctttcagctg  | atagcaagtt  | tatccttcca  | gotactcaaa  | ccagaactct  | tagagccatc  | 5280 |
| ottgacocct  | ttcctctctc  | cacactcaac  | atctatccat  | cagaaaaatt  | tggtgttctc  | 5340 |
| actttcaaaa  | tgcatacaga  | gtcagagcat  | gtctcattac  | ctccaatagc  | tacactacta  | 5400 |
| gtctgaaacaa | acatcatttc  | tcactgggtg  | tattgaacaa  | acatcatttc  | tcactgggtg  | 5460 |
| tattgatagc  | atcctaagcg  | gtcttctctg  | ttcttgggtc  | ccctatatta  | gcaacacagc  | 5520 |
| agtcagagga  | gtccttttag  | aaactcaatca | gatcatgtca  | cgtoactcct  | ctacttaaaa  | 5580 |
| tccttcaatg  | ggtccatta   | caaaaagagt  | acaaaccaga  | goccttaaac  | tggtotacaa  | 5640 |
| gttccaacat  | ttgactcctg  | ttatctctct  | gaatccttat  | tctaataatta | ctgctgttgt  | 5700 |
| ccttttgctc  | cagtcacact  | gtttgattag  | taaatattta  | ttaaacaaag  | caatcctagt  | 5760 |
| ctccaaagag  | atcatagttt  | attggaggaa  | acaagagcct  | ataaatgggt  | acacacagaa  | 5820 |
| ggtagtgttt  | atggttctcc  | ctcacctccc  | atcctaaact  | ttgacaggtg  | aaactcccct  | 5880 |
| ggatgttgaa  | ggttgaggaa  | tttgccaggg  | ttcagggttg  | tggttgaggga | ggcagggagg  | 5940 |
| aagcaaggac  | atttcaggca  | ggaagaacat  | tacatgcaaa  | gatctaaaga  | tatgaatcag  | 6000 |
| caacatatttt | atggaattac  | aagtaaaagta | gaaagttctt  | gctaaaaacat | caaaaaataa  | 6060 |
| agattttgtga | ttaggggggc  | agaatgtggg  | agggaaagag  | agatacagtt  | cacactttta  | 6120 |
| gacaggagcc  | agatcatgaa  | atgttttctc  | tttgtttgtt  | tcttctctca  | cagcttttga  | 6180 |
| tatgctcttg  | gagcaattta  | ttacccatat  | tttttaatgc  | atctcctgaa  | cagagtcata  | 6240 |
| gcaataacttg | gaaaggactc  | tgaatttctc  | gatttaaaaga | tacaaaagaa  | aaatctggag  | 6300 |



|             |            |             |             |             |             |      |
|-------------|------------|-------------|-------------|-------------|-------------|------|
| tcacaattaa  | tttgagaagg | taaaggagtg  | gggtgtgctac | tgtatcaa    | ttattttgta  | 6360 |
| caaaatcatc  | atctctagta | acattatttt  | ttctaatacta | ctgcgttttag | actacttttag | 6420 |
| taaagcttga  | tctccctgtc | tatctaaaca  | ctgattcact  | tacagcaagc  | ttcagggtag  | 6480 |
| cattgggtcat | attaataccc | acaaaatcoa  | caagggtgta  | gttgcaatg   | attttgtata  | 6540 |
| aaagggtgaac | tgagatttca | ttcagttotac | agotcttgcc  | aggcaaggca  | gcogaccaca  | 6600 |
| gggtgagctt  | ggcatctacc | gtttttoaagt | gtgacagcta  | cttttgaaat  | tacagatttg  | 6660 |
| tcaggacatg  | gaggacaaaa | ctagagcttc  | tcactactgt  | tgtgtaggaa  | attttatgctt | 6720 |
| gtcaacctgg  | cttgtaaaat | atgggttaata | taacgtaatc  | actgttagca  | agtaactgac  | 6780 |
| ttttatagacc | aatatgcctc | tcttctgaaa  | tggtcttatt  | tttaacaaat  | gtgagcaaaa  | 6840 |
| gaaaaatatt  | atgagattct | aaaaatgaag  | acataatttt  | gtagtataga  | attttcttggt | 6900 |
| ccagggaatgg | tggctcatgc | ttgtaatccc  | agcaactttg  | gaggccaagg  | tcagaggatt  | 6960 |
| gcttgagcct  | ggaagggtga | agatgcagtg  | atcoatgatt  | ataccactgc  | actccagcct  | 7020 |
| gggcaacaga  | gcaagacctc | gtotoaagaa  | aagaaaagaa  | ttttattttt  | cttttcagac  | 7080 |
| aaaaatagac  | tttaaaataa | taatggaaga  | acaaatatga  | tgatcacaa   | tatcagagta  | 7140 |
| attaactttat | gacagtcagc | aataagatto  | taattctttaa | atattcctct  | gcttaaatca  | 7200 |
| ttatatgtga  | gttttgatct | ataatatatt  | ccacacctga  | cccaaaaatt  | gaagaaggac  | 7260 |
| aaggaaaaat  | gttggttcaa | gaaacaaaga  | tgtaagtaaa  | aaggcataag  | gaaggaaaaa  | 7320 |
| aaacttttga  | agcaaatgt  | gattgaggag  | gatgagcaga  | ccaattattt  | ttggtttggt  | 7380 |
| cagcttacat  | aatgattatc | gttctttggt  | ttctcagttt  | ctagtgggt   | tcattgtttg  | 7440 |
| cttccagac   | caggatgaag | acactocagt  | ttttcttctc  | ttttgtttgc  | tggaaggcaa  | 7500 |
| tctgtgcaa   | tagctgtgag | ctgaccaaaca | tcacacattgc | aatagagaaa  | gaagaatgtc  | 7560 |
| gtttctgcat  | aagcatcaac | acacatttgg  | gtgtgtgcta  | ctgtacacac  | agggtaggta  | 7620 |
| cc          |            |             |             |             |             | 7622 |

&lt;210&gt; 2

&lt;211&gt; 6038

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 2

|             |             |            |             |             |             |      |
|-------------|-------------|------------|-------------|-------------|-------------|------|
| ggatocagaga | acatagaagg  | agcaggtaat | ttatcaaggc  | atgaacacgg  | gtgcttaatt  | 60   |
| tcttatatttg | agggcaggca  | tggtggctca | cactgttaat  | cccaacactt  | taggaagcca  | 120  |
| aggtgggtgg  | attgcttgag  | tctaggattt | tgagaccagc  | ctggccaaca  | tgggcaaatc  | 180  |
| ctgtctctac  | taaaaaatac  | aaaattaac  | agtoatgggt  | gtggtgtgco  | tttagtccca  | 240  |
| gctactctgg  | tggctgaggo  | acaagaatca | cttgaacctg  | ggaggcagag  | gttgoagtga  | 300  |
| gotgagactg  | tgccacttca  | ctocagcctg | ggtgaocagag | taagattctg  | tctcaaaaaa  | 360  |
| tatgtatata  | taacacacata | taatagatag | ataaacatat  | atacatatat  | aatatataaa  | 420  |
| tatatatat   | atatataata  | tataaacata | tataaatata  | tatatatata  | tatatatata  | 480  |
| tatatataaa  | aaacataaag  | gaataatttt | gggggaaaat  | cttcataaat  | gaaagaacaa  | 540  |
| cataggctgt  | tgagtatatg  | caacagaatt | caagagatct  | tccagcaatt  | gaagacattg  | 600  |
| gtttacoaga  | attcacaaaa  | gaagtcagct | gtgcatttaa  | agtagaatgt  | gatgagtgtt  | 660  |
| accactgagg  | taggaactgg  | gaactaagga | agcgttaagac | agaaagtgtc  | gaactgagag  | 720  |
| ttgggcattg  | gaggtctgtg  | aaggcagggt | aagtgaatgt  | ctootagaag  | ctacctttta  | 780  |
| atggagtttt  | gaagtacttg  | taggagtagc | ttagggtgaa  | agaagaggag  | aaacatgtat  | 840  |
| caggcagagg  | gactagaacc  | ttattacctt | caaggaagaa  | gcaaaaagaa  | tacatgtgac  | 900  |
| tttgagggtg  | tgggaggtgc  | tttaagccaa | tatagggtgaa | tttgacatag  | gacttcocta  | 960  |
| aataatgttc  | ggtcatattg  | taaatattga | gtgatataac  | actgtattaa  | agcccaagag  | 1020 |
| ttgcttttat  | atagaagaa   | gaaaaaagcc | caagagagtt  | ttatttctag  | agggaaatatt | 1080 |
| ttotagaaat  | aaaggaaagt  | gtatcagcca | gtttotagt   | aggaaaacag  | aatcacaco   | 1140 |
| tgatatgcaa  | aatagaggaa  | aatcagggaa | ttcattaatc  | cagagatttg  | gttgctcaag  | 1200 |
| tattagattg  | ctgaaaagcc  | agacagggaa | tatgaggcaa  | tcagagataa  | gtattagtga  | 1260 |
| caagotocat  | ttatgtgcag  | gattggaggg | acatagggtg  | ggttccocaga | agccagaagg  | 1320 |
| tgagaccacc  | tagcagaagc  | tcaaacacaa | gctgggggtt  | cctcacaana  | gctgggacca  | 1380 |
| ccaggaggag  | ctgtccaatg  | ggatctggag | ccaggggagat | catgcagtc   | ctaccaggaa  | 1440 |
| gggaagcaga  | atgtaaaagg  | tagagagaaa | tactccaact  | gcttcccttc  | attcactttc  | 1500 |
| caatctccat  | toacaaaggc  | aaaaacotgc | taatacagca  | gagtgggaaa  | agcagcctgc  | 1560 |
| caaggctcct  | tctccacaaa  | aacagagcac | aaaaccaagc  | aaaaaocagg  | aatgcatttg  | 1620 |
| atagbaaaca  | ggctatggac  | caacccaaca | taaaagaaat  | gatgagtgt   | ttcttttttc  | 1680 |
| atttgggttca | agaaaagtat  | ttoagtaact | attatgtaac  | agaaattcta  | tttatttttg  | 1740 |
| ggaattcaaa  | ggtgaataaa  | aaagaaotct | aaatttttat  | caataaata   | tttcaaaaac  | 1800 |
| ctcaatgaga  | gtaatggcat  | taactagcaa | atatgtctaat | gagatgagct  | agccataaga  | 1860 |
| ggcttagaat  | tgagagaaag  | gtctgggggc | ctcttgacag  | gccaaattca  | gagotgtttg  | 1920 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tggaatctc   | tgacctact   | gcaggtggaa  | atataaatat  | gggcatttag  | aatagtggcc  | 1980 |
| caaacttttg  | atgatttctg  | tcttggggtc  | tctccaatta  | atgggattga  | tgagaactgt  | 2040 |
| agaccactga  | ggcaccatg   | gctcaatgaa  | tagtccctg   | gctttggagt  | caaactgacc  | 2100 |
| tgaatatgaa  | cccagcttt   | gotacttaca  | ggttgcattt  | atcctcagtt  | ttctcatctt  | 2160 |
| tcaaagaaga  | acagtaactt  | ctttaaagg   | ttattgtagg  | ctgggtgcag  | tggtcaogc   | 2220 |
| ctgtaatcgc  | agcactttgg  | gaggcggagg  | ctagtggatc  | aottgagggc  | aggagtggga  | 2280 |
| aactagootg  | gccaacatgg  | tgaactctg   | tctctacaaa  | aagaaattta  | aaaaattttg  | 2340 |
| ctgggtgtgg  | tggaacacac  | ctggaattoc  | agctacotgg  | gaggcggagg  | catgagcaco  | 2400 |
| acttgagtct  | ggaaagcaga  | gggttgaggt  | gagccaagat  | tgtaccactg  | tactcaagcc  | 2460 |
| tggttgacac  | agttagacct  | tgtctaaaaa  | aaaaaaaggt  | tattgtgtta  | ttgtaaatat  | 2520 |
| tgtatatgaa  | cttctattta  | acatgtttag  | ttaaatgcot  | gtgtaattgt  | ccaatgtgct  | 2580 |
| cttctagctc  | actgcacaga  | oaaaaactgat | tcactgaaat  | catggaattg  | cagcaagaa   | 2640 |
| caaactctaat | taatgtaggt  | caaaocggag  | gactggagtt  | attattcaaa  | tcagtctccc  | 2700 |
| tgaaaaactca | gaggctaggg  | ttttatggat  | aatgttggtg  | gcaggggact  | aggggaatgg  | 2760 |
| tgctgctgat  | tggttgggga  | atgaaatagt  | aagattgtgg  | aaaactgtoc  | tcottcattg  | 2820 |
| agtctgtctc  | cgggtgtagg  | ccacacgacc  | agttgagtca  | tgaagcatgc  | gtccaagtgg  | 2880 |
| agtoagtttg  | ttgcacgaat  | gcaaaagcct  | gaaaaatgtc  | tcaaatgato  | aactgtaggc  | 2940 |
| tcacacataa  | tgatattatc  | tataggagca  | attgggggaag | taacaaatct  | tgtgacctct  | 3000 |
| ggcacacata  | actctgaact  | atgaagggat  | tataaaaaoc  | atgcotatat  | ottatoagaa  | 3060 |
| ttcaggtccc  | ccataaatoc  | taatctocaa  | gcatttccatt | tgtttagaaa  | ggccattttc  | 3120 |
| agtccttgag  | caaggagggg  | gttagtttta  | ggataggact  | attatccttg  | cttogttaaa  | 3180 |
| ctataaaacta | aattctctcc  | atgggttagct | tgccctacac  | ctaagaatga  | gtgagaaacg  | 3240 |
| ccagcctgtg  | aggctagagg  | caagatggag  | tcagccatgc  | tagattttatc | tcactgtcat  | 3300 |
| aacctttgca  | aaggcagttt  | cacotgggac  | ataggaggta  | ctcaatgaaa  | aagaagctat  | 3360 |
| taatatataa  | attttaaaaa  | tgaatttaag  | gaactaatac  | tatgtacata  | ttagtcatta  | 3420 |
| aaacaaagtg  | gttcattttac | attcacacaa  | ataaatcttg  | tgattatata  | taggtaatat  | 3480 |
| gaaaaacttt  | gttttctttc  | ataataoaag  | gtattagcaa  | tagatatagt  | aatgttagca  | 3540 |
| ttcctttgga  | aaaaatgaaa  | agatttataa  | ttttcocaaga | atcatttagta | tttttattta  | 3600 |
| atatacatata | tataaaattt  | attcattcta  | taacttggaa  | atatgotttgc | ttacraattta | 3660 |
| ctgacagatt  | tcaaaatatt  | tctatactoa  | caatattcat  | ttacataaat  | attgatttgg  | 3720 |
| taactacaat  | gtgtactgct  | atgctaagtt  | ttgtctttgt  | caaacatatt  | ttataaaatc  | 3780 |
| ataatcctag  | atgaatccaa  | cttttggtta  | cccacgtgoc  | tgaacccctg  | ctgttaacag  | 3840 |
| gcaaatgtgtg | gtaggtacag  | atctatacct  | accaotttcc  | tctacccacc  | agcatctgca  | 3900 |
| cccaccaccc  | ctcccccacc  | accatttatct | ataccaacca  | cccctcccaa  | cctaccagca  | 3960 |
| tctgcaccca  | ccacacccgc  | caccacccac  | catgtacact  | cactaacact  | tcagccatc   | 4020 |
| accatctgca  | cccatcactc  | ctcccccacc  | acaagcatct  | gcacccacca  | catttcccta  | 4080 |
| octaccagca  | tcttcaactca | ccacctctcc  | acccaccagc  | atctgcaacc  | acaacccctc  | 4140 |
| ctcacccacc  | agagtctgca  | tcacacacac  | ttgcccactc  | gctagcatct  | gcacatcaa   | 4200 |
| gotctgoott  | cttgccctaat | acgggatgag  | ctctccatgg  | ttctgcttaa  | agacaatgct  | 4260 |
| tcactcctc   | ttctataacc  | caattctctt  | tacctcttca  | agtaacactc  | agaactctct  | 4320 |
| tctcctcttg  | ataccaaactt | tttccacttt  | actcaatcat  | tcctatcaac  | atacaaacgt  | 4380 |
| gtttattttot | ccatctttaa  | agttaaaaat  | caaaaagaaa  | ttgtctgagg  | ccaggcaagg  | 4440 |
| tggtctcagc  | ctgtaatccc  | aacacttttg  | gaggccaagg  | agggttggat  | gaottaaagt  | 4500 |
| taggagttoa  | agaccagcct  | ggccaacatg  | gtgaaaccca  | totctactaa  | aaatacaaaa  | 4560 |
| attagccagg  | catggttgca  | catgctgtga  | gtctcaggta  | cttgggaggc  | tgaggccaga  | 4620 |
| gaatggcttg  | aaocggggag  | gcagaggttg  | cagttagccg  | agattgtgoc  | cttgcactcc  | 4680 |
| agcctgggtg  | acagagttag  | actccatctc  | aaaaataaaa  | aatataaata  | aaacaaaaga  | 4740 |
| aaagtatttt  | taoccaacat  | ccacattaac  | caataaccca  | tttctttatt  | gatottttgta | 4800 |
| aaaaaaagct  | cttggaaaaa  | ttgtctatat  | tcactatgac  | ttatctcctc  | caaatcactt  | 4860 |
| aaacacatac  | caatcaggtt  | tttgttttca  | tcattccaaa  | gtaactttta  | cagccaagga  | 4920 |
| oagtagcgaa  | ctttacatog  | catatgcatt  | gtgaagttct  | tgatcctcat  | cttacttaac  | 4980 |
| ctgtcagcag  | tatctgacac  | aggtgtact   | ggctcctccc  | tgagatgctc  | tctttatttg  | 5040 |
| gctttgggga  | caccatattc  | tcccccattcc | tactttctcc  | aatggccctc  | ctcagctctc  | 5100 |
| tttggaagaa  | ggaaaaagaa  | acttccattat | ctcctggatg  | tagtaaaaa   | aactcaagct  | 5160 |
| caacatgtgc  | atactgaact  | ccatttctct  | ttccaaaact  | togacattta  | cagccatccc  | 5220 |
| ctttcagctg  | atagcaagtt  | tatccttcca  | gctactcaaa  | ccagaatctt  | tagagccatc  | 5280 |
| cttgaccctt  | ttcctctctc  | cacactcaac  | atctatccat  | cagaaaaatt  | tggttggtct  | 5340 |
| actttcaaaa  | tgcatcacaga | gtcagagcat  | gtctcattac  | ctccaatagc  | taccatacta  | 5400 |
| gtctgaacaa  | acatcatttc  | tcacctgggt  | tattgaacaa  | acatcatttc  | tcacctgggt  | 5460 |
| tattgatagc  | atcctaagcg  | gtcttctgt   | ttcttggttc  | ccctatatta  | gcaaacacagc | 5520 |
| agtcagagga  | gtccttttag  | aactcaatca  | gatcatgtca  | cgtcaactct  | ctacttaaaa  | 5580 |
| tccttcaatg  | ggtcccatga  | cacaaagagt  | acaaacagaa  | gcccttaaac  | tggtctacaa  | 5640 |



|            |             |            |             |             |            |      |
|------------|-------------|------------|-------------|-------------|------------|------|
| gttccaacat | ttgactcctg  | ttatctctct | gacatcatat  | totaatatta  | ctgtgtgtgt | 5700 |
| ootttttgtc | cagtcacact  | gtttgattag | taaataattta | ttaaacaag   | caatcctagt | 5760 |
| ctccaaagag | atcatagttt  | attggaggaa | acaagagcct  | ataatgggtt  | acacacagaa | 5820 |
| ggtagtgtat | atggttctcc  | ctcacctccc | atcctaact   | ttgacaggtg  | aaactccct  | 5880 |
| ggatgttgaa | gggttgaggaa | tttgccaggg | ttcaggggtg  | tggttgaggaa | ggcaggagg  | 5940 |
| aagcaaggac | atttcaggca  | ggaagaacat | tacatgcaaa  | gatctaaga   | tatgaatcag | 6000 |
| caacatattt | atggaattac  | aagtaaagta | gaaagtto    |             |            | 6038 |

&lt;210&gt; 3

&lt;211&gt; 542

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 3

|            |            |             |            |            |             |     |
|------------|------------|-------------|------------|------------|-------------|-----|
| tcaotgttag | caagtaactg | actttataga  | ccaatatgcc | totcttctga | aatgggtctta | 60  |
| ttttaaacaa | atgtgagcaa | aagaaaatat  | ttatgagatt | ctaaaaatga | agacataatt  | 120 |
| ttgtagtata | gaattttctt | ggccagggaat | ggtgggtcat | gottgtaatc | ccagcaacttt | 180 |
| gggaggccaa | ggtcagagga | ttgcttgagc  | ctggaaggtt | gaagatgcag | tgattcatga  | 240 |
| ttataccaat | gcactccagc | ctgggcaaca  | gagcaagacc | ctgtctcaag | aaaagaaaag  | 300 |
| aattttattt | ttcttttcag | acaaaaatag  | actttaaaat | aataatggaa | gaacaaatat  | 360 |
| gatgatccca | attatcagag | taattacttt  | atgacagtoa | gcaataagat | totaatcttt  | 420 |
| aaatattcct | ctgottaaat | catttatattg | gagttttgat | ctataatata | ttccocccot  | 480 |
| gaoccaaaaa | ttgaagaagg | acaaggaaaa  | atgtttgttc | aagaaacaaa | gatgtaagta  | 540 |
| aa         |            |             |            |            |             | 542 |

&lt;210&gt; 4

&lt;211&gt; 3213

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

|             |            |             |             |             |             |      |
|-------------|------------|-------------|-------------|-------------|-------------|------|
| actaacataa  | agctgaaggt | gaataaaaaa  | atcagggtta  | gccaaacaaa  | ttttcatggt  | 60   |
| caaataccac  | ataaaaagta | aataacttta  | agttccagc   | aaaatctgaa  | ttgaacgtag  | 120  |
| acaaaatgct  | cattttctag | tgtttgacag  | acttaacagt  | ttgagccaat  | aaaastgtac  | 180  |
| tgactagata  | aactactaaa | agttgttaat  | ttttgcaatg  | tatatttctg  | aaaagaaaag  | 240  |
| ttatctatta  | tagaaattoc | tggtgocatt  | taagaacttt  | gagcatttta  | attgtttaat  | 300  |
| aataatgttt  | aattgcatca | tgaaaataat  | caataatata  | atttatttgg  | tttattttaa  | 360  |
| aaaactgatt  | ctttctgtct | ctcttatata  | tagaotgatt  | ttataactaat | gttgoottaa  | 420  |
| gatoaccaaa  | ttgtttgaag | cotaggtttc  | tgagggtatg  | aaaatgatgt  | caaaaactatt | 480  |
| tacagttcac  | acacacatto | tggttgattta | atacatcctt  | tacaagtgoa  | ggaaaggtgg  | 540  |
| aagattgatg  | atttggggga | attagagcta  | ccacacocca  | gaggggtgga  | tggtatgttg  | 600  |
| totgtttgtg  | gctgtgtgaa | tcagagagtt  | tgatttagac  | atataatttag | aaagaggaaa  | 660  |
| gatgaaccaa  | tcaaaaataa | taactataat  | gacttttcaa  | gatatagaca  | atacagttta  | 720  |
| gatataaatg  | gaacaaaaaa | aagttaaaag  | tggtggagatg | aagtctgatt  | ttttggtttt  | 780  |
| tttttttttt  | tgcttttttg | tttgtttatg  | taatcagttg  | taccagttta  | aaataatggg  | 840  |
| ttataagaca  | ctatatgcaa | gcotcatggt  | aaotccaat   | ctaaaacata  | caacaaatat  | 900  |
| acacaaaata  | aaaaggagaa | attaaaacac  | accaccagag  | aaaatcacct  | acattaaaag  | 960  |
| aaagacaaat  | aggaagaaaa | taagaaagag  | aaggccatca  | aataatcaga  | aaatgaataa  | 1020 |
| caaaatgaca  | ggaataagtc | ctcataaata  | ataacattga  | atgtaaatgg  | actaagotct  | 1080 |
| ccaatgaag   | acagggagtg | gctgaatgta  | ttttaaaaaa  | aatattcacac | cgagctgtgc  | 1140 |
| gtggtgtctc  | acacctataa | tcccagcatt  | ttgggagact  | gagccgggtg  | gatcacttga  | 1200 |
| gcccaggagt  | tcgagaccag | cctggccaac  | atggcaaaac  | cotgtctota  | ctaaaaatat  | 1260 |
| aaaaaattag  | ctgaacatgg | tggaacatgc  | ctgtggttcc  | agctactaga  | gaggctgagg  | 1320 |
| oagaagaatt  | gottgaactt | gggaggtgga  | ggttgoagtg  | agctaagatt  | gatggagcaa  | 1380 |
| ctgcacccca  | gcttaggtga | cagaataaga  | ctctgcctca  | aaaaaaaaaa  | gcaaaacaaa  | 1440 |
| acaaaacaaa  | aaaccttag  | acccaatgat  | tcattgccta  | caagaagtat  | gcttccactt  | 1500 |
| taaagacaca  | tatagactga | aggtaaaagg  | atggaaaaat  | attctatgcc  | tatggaaaca  | 1560 |
| acaaaaaaga  | agcagaagct | acattttatat | cagacaaaat  | agactgcaag  | acaaaaacta  | 1620 |
| tgaaaagaga  | gaaagaaggt | catttatatag | tgataaaggg  | gtccatttag  | caagagcatt  | 1680 |
| taacaattct  | aaatatatat | tcaccccaata | ctggagtact  | caggtatata  | aagcaaatat  | 1740 |
| tattagagcc  | aaagagagag | atagacagac  | cccatataca  | taataactgg  | agacttcaac  | 1800 |
| accccaacttt | cagcattgga | cagatcatoc  | agacagaaaa  | ttaaacaaaca | tcaaatttca  | 1860 |



|             |            |             |             |            |             |      |
|-------------|------------|-------------|-------------|------------|-------------|------|
| tctgcacocat | aggtcaaatg | gacctagtag  | atattttacag | aacatttgat | ccaacagotg  | 1920 |
| tagaataoac  | attcttctco | tcagcacatg  | gataattctc  | aaggatatac | caaattgctag | 1980 |
| gtcacaacac  | aaatctttaa | atttagaaaa  | aaagtgaat   | aatatcaaac | gtttttotctc | 2040 |
| accacagact  | aagaaaaaaa | gaagtcccaa  | ataaatacaa  | tctgagataa | aaaaggagac  | 2100 |
| gagacaacca  | ataccacaaa | aaattaaagg  | atcattagaa  | gatactatga | aaotatatgc  | 2160 |
| taataaattg  | gaaaacctga | acaaaataga  | taattcctag  | aaacatacaa | catactggtc  | 2220 |
| tgttaagggt  | ttgtattttt | tcatagtacc  | atgaagaaat  | acaagaattg | tttttagaac  | 2280 |
| oattcttgta  | tttcttcatg | gtttttgtat  | ttcttctatg  | aacctgaag  | aaatacaaaa  | 2340 |
| tgtgaacagg  | ccaataacaa | gtaattgagc  | agaagccata  | ctaaaaagta | tcccagaaaa  | 2400 |
| gaactcagga  | tctgatggot | tcaotgatga  | attttgccaa  | atatttaaaa | aaactaatacc | 2460 |
| aatccaactc  | aaattattaa | aaaaatagag  | gtggacagaa  | tctttccaaa | tgtattctat  | 2520 |
| gaggccagtg  | ttttttctga | ttgaatctcc  | cattatattt  | taatcacata | taaaaccaga  | 2580 |
| gaaagacaca  | ttaaaaagaa | agaaaactgt  | aggccaatat  | ctctgatgaa | cattgatgca  | 2640 |
| gaaatcctca  | acaacaaatt | agcaaaactga | attcaagaac  | acattaaaao | aatcattoat  | 2700 |
| oatgaocaa   | tgggaattgt | cotagagatt  | caagtgtggt  | taggtatgtg | cagatoaatg  | 2760 |
| ggtttaatgt  | tgtccaatga | acataatgtc  | ctccagctcc  | atccatgttc | ttgcaaatga  | 2820 |
| caggatctca  | ttctttttta | tggctaagta  | gtactccatt  | gtgtataagt | gccatatttt  | 2880 |
| ctttatccat  | tcatctgtta | gacaactaag  | ttgottccaa  | atcttagcta | ttgtgaatag  | 2940 |
| tgctgcaata  | aacatgggag | tgtaaatatt  | ttgttgacat  | actgatttca | tttcttttgg  | 3000 |
| ataaataccc  | agtagtgagg | ttgttgatc   | atatggggga  | aaatggagat | ggctaaccgg  | 3060 |
| ctcaaaaata  | tagttagaaa | aaatgaatat  | gatttagtat  | tccatagcac | aataggatga  | 3120 |
| ctactgttaa  | tgataattta | ttatatatta  | taaaataact  | aaaatagtat | aatgggatg   | 3180 |
| tatgtagcag  | agagaatga  | taaatgtttg  | aag         |            |             | 3213 |

&lt;210&gt; 5

&lt;211&gt; 6679

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 5

|             |            |             |            |             |             |      |
|-------------|------------|-------------|------------|-------------|-------------|------|
| gtcgcactgc  | aggtcaacgg | atcacttgag  | gacagtagtt | caagaccagc  | ctgggcagca  | 60   |
| tagggagaot  | gtctctacga | aaaatcaaaa  | aatttatggc | gggcatgggt  | gotcaogtot  | 120  |
| gtaatccctg  | aactttggga | catcaaggca  | agtggatcac | ttgagggtcag | gagttogaga  | 180  |
| ctagcctggc  | caacatgggt | aaacctatc   | tcactaaaa  | aatcaaaaaa  | ttagccaggc  | 240  |
| atggtggcag  | gcacctgtaa | tcocgggtac  | tcaggagggt | gaggcaggag  | aatcaottga  | 300  |
| accacaggag  | cgagaggtgc | agtgaagctga | gatoacacca | ctgcactcca  | gcctgggtga  | 360  |
| cagagcaaga  | ctctatctca | aaaaaaataa  | aaaaataaaa | aaattagcca  | ggcatggtag  | 420  |
| tgcaacacct  | tagtctcagc | taotcaggag  | gotgaggtgg | gaggatcact  | tgaacctggg  | 480  |
| gcagtcagg   | ctacagtgag | ccaagatcat  | gcactacac  | tcagcctgg   | gcaacagaga  | 540  |
| gagaccctgt  | ctctaaaaaa | ataataataa  | taagaaaaaa | aacagctctg  | tttatgtctc  | 600  |
| ctggtccata  | cataactact | tgtatatagt  | ttgcaaaact | aaagatccag  | atagtoaat   | 660  |
| tttttaggct  | gtgggcccga | tggctctgt   | cacaatcact | ctgccctgtc  | tttctagcac  | 720  |
| aaaagcagct  | ataaacaata | catacatgaa  | ttttttatag | acatcgagat  | ttgaatttca  | 780  |
| tatgatTTTT  | acattttata | aaataatott  | tttaaaaaat | ttcccotaac  | catttaaaag  | 840  |
| tgtaaaagcc  | ggccagcgog | ccatcgtcac  | gcctgtaatt | ccagcacttt  | gggaggctga  | 900  |
| ggtgggcaga  | tcacttgaga | tcacagttc   | gagaccagcc | tggccaacat  | agcaaaaacc  | 960  |
| catttctact  | aaaaataaaa | aaatttagctg | ggcatagtgg | tgcacacctg  | tgatcccagc  | 1020 |
| tacttgggag  | gotgaggcag | gagaatcgct  | tgaacctggg | aagcggaggt  | tgcatggagc  | 1080 |
| caacatcatg  | ccactgcact | ccagcctggg  | tgacagagtg | agacttctgc  | tcacagaaaa  | 1140 |
| aaaaaagtg   | aaaagccatt | cctaattcag  | tgtacatcag | tgtacatact  | caggtctgcg  | 1200 |
| tactcctgct  | ctgaggcaca | cctgagaagt  | agagttgott | ggtcacagga  | catacacatt  | 1260 |
| tcacacattaa | ctagacacta | ccaagtggcc  | atccaaggag | gttttttttt  | tacaactctac | 1320 |
| actcccccca  | gcaacaaatg | agagttactc  | cagatccttt | acaaagatgc  | tctaagocaa  | 1380 |
| gtaccagatg  | aaaacaggaa | gtgggagggg  | aagotgccag | ccccttctaa  | coatgaagaa  | 1440 |
| atacctggta  | gagocctctg | gatgotggaa  | ggatgaataa | cgggggtctc  | tgagcctgc   | 1500 |
| ccootgtcag  | atcactgtga | cttctgagcc  | tcagtccong | tctcagcccc  | atgtgtcatg  | 1560 |
| gccagtata   | atgagcccto | actctctgtt  | tggctcttat | tctcccoatg  | tggggctgaa  | 1620 |
| gtctggattg  | agccgttatt | caagatgtac  | agctttcttg | acaggaaaag  | agtgtcacag  | 1680 |
| aaacagcagg  | ggottggcaa | gatgatctaa  | ctgoaaatcc | taoctggctc  | agccaccagc  | 1740 |
| tagttctgtg  | atcttgaaca | agttttttca  | cttctctgag | gccatccctt  | ggctacaaca  | 1800 |
| caccagttgg  | ttgacaggat | gaaatgacga  | agtcocctac | acotgtaatc  | ccagcacttt  | 1860 |
| gggaggccaa  | ggcgggtgga | tggcttgagc  | ctgagaggtg | acagcatgoc  | ggcagtccto  | 1920 |



|            |             |            |             |             |             |      |
|------------|-------------|------------|-------------|-------------|-------------|------|
| acagccctcg | ttcgctctcg  | ggccctctcg | tgccctgggc  | ccactctcg   | tggaacttga  | 1980 |
| ggagcccttc | agccacccgc  | tgcaactgtg | gagcccttt   | ctgggctggc  | caaggccaga  | 2040 |
| gocgctccc  | tcagcttgca  | gggaggtgtg | gaggagagg   | ctcaagcagg  | aaacggggct  | 2100 |
| gocgacggcg | cttgccggcc  | agctggagtt | ccgggtgggc  | gtgggcttgg  | cgggcccccgc | 2160 |
| actcggagca | gcgggcccag  | cctgccaggc | cccgggcaat  | gagaggctta  | gcacccgggc  | 2220 |
| cagcggctgc | ggaggggtga  | ctgggtgccc | cagcagtgcc  | agcccgccgg  | cgctgtgttc  | 2280 |
| gctcgatttc | tcactgggco  | ttagcagcct | tcocgggggg  | cagggtctgg  | gacctgcagc  | 2340 |
| ccgcatgccc | tgagcctccc  | ctccatgggc | tcctgtggcg  | cccagagctc  | cccgacgagc  | 2400 |
| accacccccc | gtccacacag  | gcccagtgcc | atcgaccacg  | caagggtcga  | gaagtgcggg  | 2460 |
| cgccacggc  | cggaactggc  | aggcagctac | occtgcagcc  | ctggtgcgga  | atccactggg  | 2520 |
| tgaagccago | tggtctcctg  | agtctggtgg | agacttgagg  | aacctttatg  | tctagctcag  | 2580 |
| ggatcgtaaa | tacaccaatc  | agcaccctgt | gtctagctca  | gggtctgtga  | atgcaccaat  | 2640 |
| ccacactctg | tatctagcta  | ctctgatggg | gocctggaga  | acctttatgt  | ctagctcagg  | 2700 |
| gattgttaaa | acaccaatcg  | gcactctgtg | tctagctcaa  | ggtttgtaaa  | cacaccaatc  | 2760 |
| agcaccctgt | gtctagctca  | gggtatgtga | atgcaccaat  | cgacagtctg  | tatctggcta  | 2820 |
| ctttcatggg | catccgtgtg  | aagagaccac | caaacaggct  | ttgtgtgagc  | aataaagctt  | 2880 |
| ctatcacctg | ggtgcagggt  | ggctgagtc  | gaaaagagag  | tcagcggaagg | gagataaggg  | 2940 |
| tgggggcggt | ttataggatt  | tggttaggta | aaggaaatt   | acagtcaagg  | gggggtttgt  | 3000 |
| ctctggcggg | caggagtggt  | gggtcgcaag | gtgctcagtg  | gggggtgctt  | ttgagccagg  | 3060 |
| atgagccagg | aaaaggactt  | tcacaaggta | atgtcatcaa  | ttaggccaag  | gacccgccat  | 3120 |
| ttacacctct | tttgtgggtg  | aatgtcatca | gttaagttgg  | ggcagggcat  | attcactctc  | 3180 |
| tttgtgattc | ttcagttact  | tcaggccatc | tggtcgctata | tggtgcaagt  | acaggggatg  | 3240 |
| ogatggcttg | gcttgggctc  | agagggttga | cagctactct  | ggtggggcct  | tggaagtgt   | 3300 |
| tttgtgtcag | actctgtatc  | tagttaatct | agtggggagc  | tggaagacct  | tttgtgtcag  | 3360 |
| ctcagggatt | gtaaaacgcac | caatcagcgc | cctgtcaaaa  | cagacactc   | ggctctacca  | 3420 |
| atcagcagga | tggtgggtgg  | gccagataag | agaataaaa   | caggctgccc  | gagccagcag  | 3480 |
| tggtcaacgg | caaggtccc   | tatccacaat | atggcagctt  | tggtcttttg  | ctgtttgcga  | 3540 |
| taaatcttgc | tactgctcgc  | tttttgggtc | oacactgctt  | ttatgagctg  | taacactcac  | 3600 |
| cacgaaggtc | tgcaagttca  | ctcctgaagc | cactaaagac  | acgagccaac  | cggtgaggat  | 3660 |
| gaacaaactc | ggccgcgctg  | ccttaagagc | tataaactc   | acggogaagg  | tctgcagctt  | 3720 |
| oactcctcag | ccagcgagac  | cacgaaccca | ccagaaggaa  | gaaactgcga  | acacactctg  | 3780 |
| acatcagaag | gaacaaactc  | cagatgcacc | accttaagag  | ctgtaacact  | caactgcagg  | 3840 |
| gtccgcggct | tccttcttga  | agtcagtgag | accaaagcact | caccagtttc  | ggacacaagc  | 3900 |
| ccaggagttt | gagatcagcc  | tggtgcaaat | gatgaatgc   | cctctctgca  | aaaaaaaaaa  | 3960 |
| aaattacaaa | aatggcgga   | gcattggtgt | ccgtgcctgt  | gggtccagct  | acgcgggagg  | 4020 |
| ctaaagtggg | aggatcgctt  | gagcctggga | ggtgaagact  | gcagtgaagt  | gtgattgtac  | 4080 |
| caacagccct | taggctgggg  | gaagagctga | gacctgttt   | ccctcccgca  | aaaaaattga  | 4140 |
| caaaagtgtg | ataagaggtg  | cotgatattg | ctaggccgag  | tggtcatgc   | ctgtaatccc  | 4200 |
| agcaacttgg | gaagccgagg  | ogggcgggtc | acctaaggtc  | aggagtgtga  | gaacagctgt  | 4260 |
| gccaacatgg | agaaagccca  | totcttctaa | aaatacaaaa  | ttagccgggt  | gtgggggag   | 4320 |
| tggtggagca | tgccgtgaat  | cccagctact | caggaggtcg  | aggcaggaga  | atcacttgaa  | 4380 |
| ccoagggagg | ggcggttgca  | gtgagccgag | atcgtgccat  | tgcaactccac | ccactccagc  | 4440 |
| ctgggcaaca | agagccaaac  | tctgtcttaa | aaaaaaaaaa  | aaaaagtgc   | tgacatataa  | 4500 |
| gaggtgtgca | atgcaatagt  | tgccaggcaa | catgtttaag  | aatgtggagc  | toctgccttc  | 4560 |
| catgctcctg | ttaaaaaacc  | acctcaagg  | ccaggtgcag  | tggtcatgc   | ctataatccc  | 4620 |
| agcaacttgg | gagggccagg  | oggtgggac  | acctgaggtc  | aggagtgcga  | gaccagctgt  | 4680 |
| accaccaaca | tggtgaaatc  | ccactctac  | taaaaaatac  | aaattagatg  | agcatgggtg  | 4740 |
| tgcatgctgt | taatccaccc  | tacttgggag | gctgagggag  | gaaaatcact  | agaaccaggg  | 4800 |
| aggcggaggt | tgtagtgagc  | cgagatcggt | ccattgcact  | ccagcctgag  | caatgagcga  | 4860 |
| aactccatct | caaaaaaaca  | acaacaaaaa | ccactctctc  | actcccaggg  | agctgggtac  | 4920 |
| agagctgggc | caatcagtg   | caaggtgctg | agccacagag  | ctaaggcgga  | gctgacggac  | 4980 |
| cgccggaccg | ataacagtg   | gtgagatcag | tggtgtgagc  | cagacgtccc  | tgccattggt  | 5040 |
| gaccaccagg | gggcccccaa  | gcacagagag | tggtcccatc  | cagtcaccac  | atccactctc  | 5100 |
| catccagaga | tggtgtgttc  | ttggcagcgt | gggttaaat   | aggacagaa   | gtgacagtct  | 5160 |
| tggtgtggtg | cagtcagact  | gcccaggcca | ggccttgggt  | octgtagaaa  | acgttcaggc  | 5220 |
| ctaggccggg | caoggtggct  | cacgcctgta | atccagagac  | tttgggagga  | cgaggccgggt | 5280 |
| ggatccagag | gtcaggagat  | cgtgacatc  | ctggttaaca  | cggtgaaac   | cogtctctac  | 5340 |
| taaaaataca | aaaaattggc  | cggtcatggt | ggcgggcacc  | tgtagttcca  | gtactcggg   | 5400 |
| aggctgaggg | aggagaatgg  | cgtgaacccg | agaggcagag  | tttgagtgga  | gocagagatc  | 5460 |
| cgccactgca | ctccagcctg  | ggcgacagag | caagactcca  | tctggaaaag  | aaaaagaaaa  | 5520 |
| cggtcaggtc | tgccagagc   | gcccagcctg | taattctgtc  | acottaacatg | accttgggca  | 5580 |
| aggcaactcc | ttccctggcc  | cagttcacgg | ggttggaaac  | gaactcaagg  | tccttccag   | 5640 |



|             |             |             |             |            |            |      |
|-------------|-------------|-------------|-------------|------------|------------|------|
| oattaacgct  | gcatggttct  | aagatgagaa  | gatggggcag  | tttccctct  | ctcaccacag | 5700 |
| cccgtgtcca  | cttcaagggtg | aatgaccagg  | gaagtacagt  | gtcccaatcc | cgcagttcca | 5760 |
| aagcccttgg  | ggaccctact  | gtcagggtcg  | tgcacagga   | ggtgaaggto | aggtgagcca | 5820 |
| atcgctctga  | aggggtcttgc | ctcattcggg  | acagacatcc  | ggtttcctct | ggctctacog | 5880 |
| ggattctagg  | ggcttttagcc | gaatgagtoa  | tggggggcgg  | gggggtttct | gggggagtto | 5940 |
| ccagctaate  | aaactggggac | aggacagcct  | ggaactttcg  | atggtgccta | tccaagtgtg | 6000 |
| gggtggggac  | agcagccaag  | acccaatgtc  | ottatctcag  | gtagggggtc | aggaggtctc | 6060 |
| ccagacaggg  | agcctccgga  | gagtttgggg  | gtaggaaatgg | gagcaaccag | gcttcttttt | 6120 |
| ttctctctta  | gaatttgggg  | gcttggggga  | caggcttgag  | aatcccaaag | gagaggggca | 6180 |
| aaggacactc  | ccccacaagt  | ctgccaagagc | gagagagggga | gaccccgact | cagctggcac | 6240 |
| ttccccacag  | gcctctgocg  | cttccaggog  | tctatcagcg  | gctcagcctt | tgttcagctg | 6300 |
| ttctgttcaa  | acactctggg  | gccattcagg  | cctgggtggg  | gcagcgggag | gaaggggagt | 6360 |
| tgaggggggg  | aaggcgacgt  | caaaggagga  | tcaagatctc  | cacaatttca | caaaactttc | 6420 |
| gcaaacagct  | ttttgttcca  | acccctctgc  | attgtcttgg  | acaccaaatt | tgcataaatc | 6480 |
| ctggggaagt  | attactaago  | cttagtcgtg  | gcccaggtta  | atttctctcc | aggcctccat | 6540 |
| gggggttatgt | ataaagggcc  | ccctagagct  | ggggcccaaa  | acagcccgga | gcctgcagcc | 6600 |
| cagccccacc  | cagaccatg   | gctggacctg  | ccaccacagag | cccatgaag  | ctgatgggtg | 6660 |
| agtgtctctgg | cccaggatg   |             |             |            |            | 6679 |

&lt;210&gt; 6

&lt;211&gt; 6235

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 6

|             |             |             |             |             |            |      |
|-------------|-------------|-------------|-------------|-------------|------------|------|
| gatcaactga  | ggacagtagt  | tcaagaccag  | cctggggcagc | atagggagac  | tgtctctacg | 60   |
| aaaaatcaaa  | aaattatggc  | cgggcatggt  | ggctcagctc  | tgtaatccct  | gaacttttgg | 120  |
| acatcaaggg  | aagtggatca  | cttgagggtca | ggagttcag   | actagcctgg  | ccaacatggt | 180  |
| gaaacccat   | ctccactaaa  | aaatacaaaa  | attagccagg  | catggtggga  | ggcactgta  | 240  |
| atccccggta  | ctcaggaggg  | tgaggcagga  | gaatcaactg  | aaacccaggag | goggaggttg | 300  |
| oagtgagctg  | agatcacacc  | actgcactcc  | agcctgggtg  | acagagcaag  | actctatctc | 360  |
| aaaaaaaata  | aaaaataaaa  | aaaattagcc  | aggaatggta  | gtgcacacot  | ctagtctcag | 420  |
| ctaactcagga | ggctgaggtg  | ggaggatcac  | ttgaacctgg  | ggcagtcagg  | gctacagtga | 480  |
| gccaagatca  | tgccactaca  | ctccagcctg  | ggcaacagag  | agagaccctg  | tctctaaaaa | 540  |
| aataataata  | ataaagaaaa  | aaacagctct  | gtttatgtot  | cctgggtccat | acatactact | 600  |
| atgtatatag  | tttgcaaaact | caaagatcca  | gatagtoaat  | tttttaggct  | tgtgggcogt | 660  |
| atggtctctg  | tcacaatcac  | tctgcctgtg  | ctttctagca  | caaaagcagc  | tataaacaat | 720  |
| acatacatga  | attttttata  | gaactcgaga  | tttgaatttc  | atatgatttt  | tacattttat | 780  |
| aaaataaatct | ttttaaaaaat | tttccoctaa  | ccatttaaaa  | gtgtaaaagc  | cggccagcgc | 840  |
| gccatcgtca  | cgcctgtaat  | tccagcactt  | tgggaggtcg  | aggtggggcag | atcacttgag | 900  |
| atcaacagtt  | cgagaccagc  | ctggccaaca  | tagcaaaacc  | ccatttctac  | taaaaataaa | 960  |
| aaaatttagct | gggcatagtg  | gtgcacacot  | gtgatcccag  | ctacttggga  | ggotgaggca | 1020 |
| ggagaatcgc  | ttgaacctgg  | gaagcggagg  | ttgcagttag  | ccaacatcat  | gcaactgcac | 1080 |
| tccagcctgg  | gtgacagagt  | gagaactcgt  | otcaacgaaa  | aaaaaaagtg  | taaaagccat | 1140 |
| tactaatcca  | gtgtacatca  | gtgtacatac  | toaggtctgc  | gtactoctgc  | tctgaggcat | 1200 |
| acotgagaag  | tagagttagct | tgggtcacagg | acatacacat  | ttccacatta  | actagacact | 1260 |
| accaagttgc  | catccaagga  | gggttttttt  | ttacaatcta  | cactcccccc  | agcaacaat  | 1320 |
| gagagttact  | ccagatcctt  | taaaaagatg  | ctctaagccc  | agtaccagat  | gaaaacagga | 1380 |
| agtgaggagg  | gaagctgcca  | gccccttcta  | accatgaaga  | aatacctggt  | agagccttct | 1440 |
| ggatgctgga  | aggatgaata  | acgggggtot  | ctggagcctg  | ccccctgtca  | gatcactgtg | 1500 |
| acttctgagc  | ctccagtcga  | gtctcagccc  | catgtgtcat  | ggccagtgat  | aatgagccct | 1560 |
| cactctctgt  | ttggtcttta  | ttctccccat  | gtgggggtga  | agtctggatt  | gagccgttat | 1620 |
| tcaagatgta  | cagctttctt  | gaacggaaag  | tagtgtcaca  | gaaacagcag  | ggccttggca | 1680 |
| agatgatcta  | actgcaaatc  | ctacctggct  | cagccaccag  | ctagtctctg  | gatcttgaac | 1740 |
| aagttttttc  | acttctctga  | ggccatccct  | tggctaaaac  | acaccagttg  | gttgacagga | 1800 |
| tgaatgacg   | aagtccctta  | cacctgtaat  | ccagcactt   | tgggagggca  | aggcgggtgg | 1860 |
| atggcttgag  | cctgagaggt  | gacagpatgc  | cggcagtcct  | cacagccctc  | gttcgctctc | 1920 |
| ggcgcctctc  | ctgcctgggc  | tcccacttcg  | gtggcacttg  | aggagccctt  | cagcccaacc | 1980 |
| ctgcactgtg  | ggagcccctt  | tctggggtgg  | ccaaggccag  | agccggctcc  | ctcagcttgc | 2040 |
| agggaggtgt  | ggagggagag  | gctcaagcag  | gaacgggggc  | tgcgcacggc  | gcttgccggc | 2100 |
| cagctggagt  | tcgggtggg   | cgtgggcttg  | gcgggcccgc  | caotggagc   | agcggggcag | 2160 |
| ccctgcccagg | cccogggcaa  | tgagaggctt  | agcaocccgg  | ccaggggtgt  | cggaggggtg | 2220 |



|            |             |            |             |            |            |      |
|------------|-------------|------------|-------------|------------|------------|------|
| caggacagcc | tgggaactttc | gatggtgcct | atcgaagtgt  | ggggtgggca | cagcagccaa | 6000 |
| gaaccaatgt | ccttatotca  | ggtaggggct | caggagggtct | cccagacagg | cagcctccgg | 6060 |
| agagtttggg | ggttaggaatg | ggagcaacca | ggcttctttt  | tttctctctt | agaatttggg | 6120 |
| ggcttggggg | acaggccttga | gaatcccaaa | ggagaggggc  | aaaggacact | ccccacaag  | 6180 |
| tctgccagag | cgagagaggg  | agaccccgac | tcagctgona  | cttccccaca | ggcct      | 6235 |

<210> 7  
 <211> 278  
 <212> DNA  
 <213> Homo sapiens

|            |            |             |            |            |             |     |
|------------|------------|-------------|------------|------------|-------------|-----|
| <400> 7    |            |             |            |            |             |     |
| aagcttttat | aggtgtaaat | tttccaactta | gtactgottt | tgtaatgttg | tottttttatt | 60  |
| ttcatttato | tcaagatggt | ttctaatttc  | tcttgacttc | cttcttaaat | tottacotca  | 120 |
| tgtagacata | catttttggc | octatgcatt  | gggatgaaaa | accagactaa | tttaotttgt  | 180 |
| acaaaaagaa | aaatgagaaa | gaaatatatt  | tggtcttgtg | agcactatat | ggaaataott  | 240 |
| tatatccat  | ttgtttcato | atatccat    | atccctt    |            |             | 278 |

<210> 8  
 <211> 73  
 <212> DNA  
 <213> Homo sapiens

|            |           |          |            |             |            |    |
|------------|-----------|----------|------------|-------------|------------|----|
| <400> 8    |           |          |            |             |            |    |
| cattggatc  | ttcatccct | gctgtgat | tatgaatgtc | tgccctatata | aatattcact | 60 |
| attccataac | aca       |          |            |             |            | 73 |

<210> 9  
 <211> 3033  
 <212> DNA  
 <213> Homo sapiens

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| <400> 9     |             |             |             |             |             |      |
| actaacataa  | agctgaaggt  | gaataaaaaa  | atcagggtta  | gccaaacaaa  | ttttcatggc  | 60   |
| caaatapoc   | ataaaaaagta | aatatactta  | agttcccagc  | aaaatctgaa  | ttgaacgtag  | 120  |
| acaaaatgct  | cattttctcag | tgtttgacag  | acttaacagt  | ttgagocaa   | aaaaatgtac  | 180  |
| tgactagata  | aactactaaa  | agttgttaat  | ttttgcaatg  | tatatattctg | aaaagaaagt  | 240  |
| ttatctatta  | tagaaaattcc | tggtgccatt  | taagaaacttt | gagcatttta  | attgtttta   | 300  |
| aatatagttt  | aattgcatca  | tgaaaaataat | caataataca  | atttattttg  | tttattttaa  | 360  |
| aaaactgatt  | ctttctgtctc | tctctatata  | tagactgatt  | ttataactaat | gttgccctaaa | 420  |
| gatcacocaa  | ttgtttgaag  | cctagggtt   | tgagggatgg  | aaaatgatgt  | cacaactatt  | 480  |
| tacagttoac  | acacacattc  | tggggattta  | atacatcctt  | tacaagtgc   | ggaaagggtg  | 540  |
| aagattgatg  | atttggggga  | attagagcta  | ccacacccca  | gagggtggta  | tggtatgttg  | 600  |
| tctgttgtga  | gctgtgtgaa  | tcagagagtt  | tgatttagac  | atatatttag  | aaagaggaaa  | 660  |
| gatgaaccaa  | taaaaaataa  | taactataat  | gacttttcaa  | gatatagaca  | atacagttaa  | 720  |
| gatataaatg  | gaacacaaaa  | aaagttaaaag | tggggagatg  | aagtctgatt  | ttttgggttt  | 780  |
| tttttttttt  | tgctttttttg | tttggtttatg | taatcagttg  | taocagttta  | aaataatggg  | 840  |
| ttataagaca  | ctatatgcaa  | gootcatggt  | aacctocaa   | ctaaaacata  | caacaaatac  | 900  |
| acacaaaata  | aaaaggagaa  | attaaaacac  | accaccagag  | aaaatcaoct  | acattaaaag  | 960  |
| aaagacaaat  | aggagagaaa  | taagaaaagag | aaggcoatca  | ataatcaga   | aatgaataa   | 1020 |
| caaaatgaca  | ggaataagtc  | ctcataaata  | ataacattga  | atgtaaatgg  | actaagctot  | 1080 |
| ccaatgaaag  | acagggagtg  | gctgaatgta  | ttttaaaaaa  | aatattacac  | cgagctgtgc  | 1140 |
| gtggtgtctc  | acacctataa  | tcocagcatt  | ttgggagact  | gagocgggtg  | gatoacttga  | 1200 |
| gcccaggagt  | tcgagaccag  | cctggccaac  | atggcaaac   | ootgtctcta  | ctaaaaatac  | 1260 |
| aaaaaattag  | ctgaacatgg  | tggaacatgc  | ctgtgggtcc  | agctactaga  | gagggtgagg  | 1320 |
| cagaagaatt  | gcttgaactt  | gggaggtgga  | gggtgcagtg  | agctaagatt  | gatggagoca  | 1380 |
| ctgcacccca  | gcctagggtga | cagaataaga  | ctctgcctca  | aaaaaaaaaa  | gcaaaacaaa  | 1440 |
| acaaaacaaa  | aaacccttag  | acccaatgat  | tcattgccta  | caagaagtat  | gottcacctt  | 1500 |
| taaaagacaca | tatagactga  | aggtaaaagg  | atggaaaaat  | attctatgoc  | tatggaaaaca | 1560 |
| aaacaaaaga  | agcagaagct  | acatttatat  | cagacaaaat  | agactgcaag  | acaaaaacta  | 1620 |
| tgaaaagaga  | gaaagaaggt  | catttatatag | tgataaagg   | gtocatttag  | caagagcatt  | 1680 |
| taacaattct  | aaatatatat  | tcacccaata  | ctggagtagt  | cagggtatata | aagcaaatat  | 1740 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| actgggtgco  | ccagcagtg   | cagcccgccg  | gcgctgtgct  | ogctcgattt  | otcactgggc  | 2280 |
| cttagcagcc  | ttcccgccgg  | gcagggtctg  | ggacotgcag  | cccgcatgc   | ctgagcotcc  | 2340 |
| octccatggg  | ctcctgtgog  | gcccgagcct  | ccccgacgag  | caaccacccc  | tgctccacag  | 2400 |
| ogccocagtc  | catcgaccac  | gaaagggtcg  | agaagtgcgg  | gcgcacggca  | ccgggactgg  | 2460 |
| caggcagcta  | cccctgcagc  | cctggtgogg  | aatccactgg  | gtgaagccag  | ctgggtcctt  | 2520 |
| gagtcctggg  | gagacttgga  | gaacotttat  | gtctagotca  | gggatcgtaa  | atacacaat   | 2580 |
| cagcaccctg  | tgtctagotc  | agggtctgtg  | aatgcaccaa  | tccacactot  | gtatctagct  | 2640 |
| actctgatgg  | ggccttgagg  | aaottttatg  | tctagotcag  | ggattgtaaa  | tacaccaatc  | 2700 |
| ggcactctgt  | atctagctca  | aggtttgtaa  | acacaccaat  | cagoaccctg  | tgtctagotc  | 2760 |
| agggtatgtg  | aatgcaccaa  | tcgacagctc  | gtatctggct  | actttcatgg  | gcacccgtgt  | 2820 |
| gaagagacca  | ccaaaacagg  | tttgtgtgag  | caataaagct  | tctatcacct  | gggtgcaggt  | 2880 |
| gggctgagtc  | cgaagagaga  | gtcagcgaa   | ggagataagg  | gtggggccgt  | tttataggat  | 2940 |
| ttgggttaggt | aaaggaaaat  | tacagtcaaa  | gggggtttgt  | tctctggggg  | gcaggagtgg  | 3000 |
| ggggctcgaa  | ggtgctcagt  | gggggtgott  | tttgagccag  | gatgagccag  | gaaaaggact  | 3060 |
| ttcacaaggt  | aatgtcatca  | attaaggcaa  | ggaccccgca  | tttacacctc  | ttttgtgggt  | 3120 |
| gaatgtcatc  | agtttaagtt  | gggcaggcca  | tattcaacttc | ttttgtgatt  | ottcagttac  | 3180 |
| ttcaggccat  | ctgggctgat  | atgtgcaagt  | tacaggggat  | gcgatggctt  | ggottggggt  | 3240 |
| cagaggottg  | acagctactc  | tggtggggcc  | ttggagaatg  | tttgtgtoga  | oactotgtat  | 3300 |
| ctagttaato  | tatgtgggac  | gtggagaaac  | tttgtgtcta  | gtcaggggat  | tgtaaacgca  | 3360 |
| ccaatcagcg  | ccctgtcaaa  | acagaccact  | oggtctatcc  | aatcagcagg  | atgtgggtgg  | 3420 |
| ggcagataaa  | gagaataaaa  | gcaggtgccc  | cgagocagca  | gtggaaagcc  | gcacaggtcc  | 3480 |
| ctatccacaa  | tatggcagot  | ttgttctttt  | gctgttttgg  | ataaatcttg  | ctactgtctg  | 3540 |
| ctttttgggt  | ccacactgot  | tttatgagct  | gtaacactca  | ccacgaaggt  | ctgcagcttc  | 3600 |
| actcotgaag  | ccactaagac  | cacgagccca  | ccgggaggaa  | tgaaacaaot  | ccggccggct  | 3660 |
| goottaagag  | ctataacact  | cccgoggaag  | gtctcagctc  | tcaactctca  | gccagcgaga  | 3720 |
| ccacgaaccc  | accagaagga  | agaaactcg   | aacacatctg  | aacatcagaa  | ggaacaaact  | 3780 |
| ccagatgcac  | cacotaaaga  | gctgtaacac  | tcactgogag  | ggtcocgggc  | ttccttcttg  | 3840 |
| aaagtgcagc  | gaccaagcac  | tcaccagttt  | cggacacaag  | cccaggagtt  | tgagatcagc  | 3900 |
| ctgggcaaca  | tgatgaaatg  | ccctctctgc  | aaaaaasaaa  | aaaaatcaaa  | aaattggcgg  | 3960 |
| agcagtggtg  | tcogtgccog  | tggtcccagc  | tacgggggag  | gctaaagtgg  | gaggatcgct  | 4020 |
| tgagcctggg  | aggtgaagac  | tcgagtgcgc  | tgtgattgta  | ccacagccct  | ctagggtggg  | 4080 |
| ggacagactg  | agaccctgtt  | tcocctccgc  | aaaaaaattg  | acaaaagtgt  | aataagaggt  | 4140 |
| gcctgatatg  | gctaggogca  | gtggtcatg   | cotgtaatcc  | cagcactttg  | ggaagccgag  | 4200 |
| gcgggcgggt  | cacctaaagt  | caggagtgtg  | agaccagcct  | ggccacacat  | gagaaagccc  | 4260 |
| atctctctta  | aaaatacaaa  | attagccggc  | tggtggggca  | gtggtgggag  | atgctgttaa  | 4320 |
| tcocagctac  | tcaggaggct  | gaggcaggag  | aatcacttga  | acccaggagg  | cggggttgc   | 4380 |
| agtgcagcca  | gatcgtgcca  | ttgcactcca  | cccactccag  | octgggcaac  | aagagccaaa  | 4440 |
| ctctgtctta  | aaaaaasaaa  | aaaaaagtgc  | ctgacatata  | agaggtgtgc  | aatgcaatag  | 4500 |
| ttgccaggca  | acatgtttaa  | gaatgtggag  | ctcctgcctt  | coatggctct  | gttaaaaaac  | 4560 |
| cacootcaag  | gccagggtgca | gtggtctcat  | cctataatoc  | cagcactttg  | ggaggccgag  | 4620 |
| gcgggtggat  | cacotgaggt  | caggagtctg  | agaccagcct  | gacccaacac  | atgggtgaaat | 4680 |
| cccactctta  | ctaaaaatag  | aaaatttagat | gagcactggt  | gtgoatgctt  | gtaatcccac  | 4740 |
| ctacttgagg  | ggctgagggc  | ggaaaatcac  | tagaacacag  | gaggcggagg  | ttgtagttag  | 4800 |
| ccgagatcgt  | gccattgcac  | tcacgcctga  | gcaatgagcg  | aaactccatc  | tcaaaaaaac  | 4860 |
| aaacaacaaa  | acccactctc  | tactcccagg  | gagctgggta  | cagagctggg  | ccacatcagt  | 4920 |
| gcaagggtgt  | gagccacaga  | gctaaggcgg  | agctgcaggc  | ccgcggacca  | gataaacagt  | 4980 |
| tgtgagatca  | gtgtgtgaga  | tcagacgtoc  | ctgocattgg  | tgacccaccag | ggggcccccac | 5040 |
| agcaccagag  | atggccccat  | ccagtcacca  | catccaottc  | tcacccagag  | atgtctgttt  | 5100 |
| ottggoacgc  | tggggtaaat  | taggaacagaa | gggtgacagtc | ttgggtgttg  | tcagtccagac | 5160 |
| tgcccagggc  | agggccttgg  | gcotgtagaa  | aacgttcagg  | cctaggcccg  | gcacgggtgg  | 5220 |
| tcacgcctgt  | aatccagcca  | ctttgggagg  | ccgaggccgg  | tggtcacga   | ggtcaggaga  | 5280 |
| tcgtgaccat  | cctggctaao  | acgggtgaaac | ccgtctctca  | ctaaaaatac  | aaaaaattgg  | 5340 |
| ccgggcatgg  | tgggggggc   | ctgtagtctc  | agctactcgg  | gaggotgagg  | caggagaatg  | 5400 |
| gcgtgaaccc  | gagaggcaga  | gtttgcagtg  | agccagagtc  | gcgccactgc  | actccagcct  | 5460 |
| ggggcagaga  | gcaagactcc  | atctggaaaa  | gaaaaagaaa  | aogttcaggt  | ctgagccaga  | 5520 |
| ggcccagggt  | gtaattctgt  | cacttaaat   | gacottgggc  | aaggcacttc  | cttccctggc  | 5580 |
| ccagttcacg  | gggttggaat  | cgactccaa   | gtcccttcca  | gcattaaagc  | tgoatgggtc  | 5640 |
| taagatgaga  | agatggggca  | gtttccctcc  | ctccacccca  | gcccgtgtcc  | acttcaaggt  | 5700 |
| gaatgaccag  | ggaagtacg   | tgtcccaatc  | ccgoagttcc  | aaagcccttg  | gggaccctac  | 5760 |
| tgtcagggtc  | gtgcacgagg  | aggtgaaggt  | caggtgagcc  | aatcgocctg  | aagggtcttg  | 5820 |
| cctcattcgg  | gacagacatc  | oggtttctcc  | tggtctatcc  | gggattctag  | gggcttttag  | 5880 |
| ogaatgagtc  | atggggggcg  | gggggggttc  | tgggggaggt  | cccagctaat  | caacttggga  | 5940 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tattagagcc  | aaagagagag  | atagacagac  | ccccatacaa  | taataactgg  | agacttoaac  | 1800 |
| accccacttt  | cagcacttga  | cagatcatcc  | agacagaaaa  | ttaacaaaca  | tcaaatctca  | 1860 |
| tctgoaccat  | aggtcaaattg | gacctagtag  | atattttacag | aacattttgat | ccaacagctg  | 1920 |
| tagaatacac  | attctttotcc | tcagcacatg  | gataattctc  | aaggatatac  | caaattgctag | 1980 |
| gtcacaaaaac | aaatctttaa  | atthagaaaa  | aaagtgaatt  | aatatcaaac  | gtttttotctc | 2040 |
| accacagact  | aagaaaaaaa  | gaagtcccaa  | ataaatataa  | tctgagataa  | aaaaggagac  | 2100 |
| gagacaacca  | ataccacaaa  | saattaaagg  | atoattagaa  | gatactatga  | aactatatgc  | 2160 |
| taataaattg  | gaaaacctga  | acaaaataga  | taattctctag | aaacatacaa  | catactggtc  | 2220 |
| tgttoagggt  | ttgtattttt  | tcatagtacc  | atgaagaaat  | acaagaattg  | tttctagaac  | 2280 |
| cattcttctga | tttcttctatg | gtttttgtat  | ttcttctatg  | aaacctgaag  | aaatacaaaa  | 2340 |
| tgtgaacagg  | coaataacaa  | gtaattgagc  | agaagccata  | ctaaaaagta  | tcccagaaaa  | 2400 |
| gaactcagga  | tctgatggct  | tcactgatga  | atthttgcaa  | atatttaaaa  | aaotaatacc  | 2460 |
| aatccaaact  | aaattattaa  | aaaaatagag  | gtggacagaa  | tottttccaaa | tgtattctat  | 2520 |
| gaggccagtg  | ttttttctga  | ttgaatctcc  | catttatatt  | taatacacata | taaaaccaga  | 2580 |
| gaaagacaca  | ttaaaaagaa  | agaaaactgt  | aggccaatat  | ctctgatgaa  | cattgatgca  | 2640 |
| gaaatcctca  | acaacaaatt  | agcaaaactga | attcaagaa   | acattaaaac  | aatcattcat  | 2700 |
| catgaccaag  | tggaatttgt  | cctagagatt  | caagtgttgt  | taggtatgtg  | cagatcaatg  | 2760 |
| ggtttaattgt | tgtccaatga  | acataatgtc  | ctccagctcc  | atccatgttc  | ttgcaaatga  | 2820 |
| caggatctca  | ttctttttta  | tggttaagta  | gtactocatt  | gtgtataaag  | gocatatttt  | 2880 |
| otttatccat  | tcatctgtta  | gacacctaag  | ttgottcoaa  | atotttagota | ttgtgaatag  | 2940 |
| tgotgcaata  | aacatgggag  | tgtaaatatt  | ttgttgacat  | actgatttca  | tttcttttgg  | 3000 |
| ataaataccc  | agtagtggga  | ttgctgggac  | ata         |             |             | 3033 |



## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

Int. l. Application No  
PCT/US 01/07870

|  |  |  |
|--|--|--|
| A. CLASSIFICATION OF SUBJECT MATTER<br>IPC 7 C12N15/63   |  |  |
| According to International Patent Classification (IPC) or to both national classification and IPC  |  |  |
| B. FIELDS SEARCHED<br>Minimum documentation searched (classification system followed by classification symbols)<br>IPC 7 C12N  |  |  |
| Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  |  |  |
| Electronic data base consulted during the international search (name of data base and, where practical, search terms used)<br>EPO-Internal, BIOSIS, WPI Data   |  |  |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT   |  |  |
| Category *   | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.  |
| A  | WO 00 12716 A (PIONEER HI BRED INT)<br>9 March 2000 (2000-03-09)<br>abstract<br>page 2, line 14 - line 24<br>page 3, line 10 - line 15   | 1-8  |
| A  | TSUKAMOTO Y ET AL: "Hdf1, a yeast Ku-protein homologue, is involved in illegitimate recombination, but not in homologous recombination"<br>NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB,<br>vol. 24, no. 11, 1996, pages 2067-2072,<br>XP002170162<br>ISSN: 0305-1048<br>the whole document | 1-9  |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.  |  |  |
| * Special categories of cited documents :<br>"A" document defining the general state of the art which is not considered to be of particular relevance<br>"E" earlier document but published on or after the international filing date<br>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)<br>"O" document referring to an oral disclosure, use, exhibition or other means<br>"P" document published prior to the international filing date but later than the priority date claimed<br>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention<br>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone<br>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art<br>"Z" document member of the same patent family |  |  |
| Date of the actual completion of the international search<br>27 September 2001   |  | Date of mailing of the international search report<br>12/10/2001 |
| Name and mailing address of the ISA<br>European Patent Office, P.B. 5818 Palenstein 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040 Tx. 31 651 epo nl<br>Fax. (+31-70) 340-3016   |  | Authorized officer<br>Panzica, G                                 |



## INTERNATIONAL SEARCH REPORT

Int. Appl. No.

PCT/US 01/07870

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT |  |                       |
|--|--|-----------------------|
| Category *   | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
| A  | PETRINI J H J: "The Mammalian Mre11-Rad50-Nbs1 Protein Complex: Integration of Functions in the Cellular DNA-Damage Response"<br>AMERICAN JOURNAL OF HUMAN GENETICS,<br>UNIVERSITY OF CHICAGO PRESS, CHICAGO,, US,<br>vol. 64, 6 April 1999 (1999-04-06), pages<br>1264-1269, XP002150324<br>ISSN: 0002-9297<br>the whole document | 1-9                   |
| A  | CRITCHLOW S E ET AL: "MAMMALIAN DNA DOUBLE-STRAND BREAK REPAIR PROTEIN XRCC4 INTERACTS WITH DNA LIGASE IV"<br>CURRENT BIOLOGY, CURRENT SCIENCE,, GB,<br>vol. 7, no. 8, 1 August 1997 (1997-08-01),<br>pages 588-598, XP002065034<br>ISSN: 0960-9822<br>the whole document  | 1-9                   |
| A  | TOMKINSON A E ET AL: "STRUCTURE AND FUNCTION OF MAMMALIAN DNA LIGASES"<br>MUTATION RESEARCH, AMSTERDAM, NL,<br>vol. 407, no. 1, 1998, pages 1-9,<br>XP002065035<br>ISSN: 0027-5107<br>the whole document   | 1-9                   |
| A  | WO 98 30902 A (CRITCHLOW SUSAN ELIZABETH ;CANCER RES CAMPAIGN TECH (GB); JACKSON)<br>16 July 1998 (1998-07-16)<br>abstract<br>claims   | 1-9                   |
| A  | WO 99 25829 A (NANDABALAN KRISHNAN ;YANG MEIJIA (US); CURAGEN CORP (US); SCHULZ V)<br>27 May 1999 (1999-05-27)   |                       |
| A  | WO 98 30903 A (CANCER RES CAMPAIGN TECH ;DOWNS JESSICA ANNE (GB); JACKSON STEPHEN)<br>16 July 1998 (1998-07-16)<br>abstract<br>disclosure<br>claims  | 1-9                   |
| A  | US 5 641 670 A (HEARTLEIN MICHAEL W ET AL) 24 June 1997 (1997-06-24)<br>cited in the application<br>the whole document   | 1,9,10,<br>26,27,75   |
| A  | US 5 272 071 A (CHAPPEL SCOTT C)<br>21 December 1993 (1993-12-21)<br>cited in the application<br>the whole document  | 1,9,10,<br>26,27,75   |

-/--



## INTERNATIONAL SEARCH REPORT

Int. l. Application No.

PCT/US 01/07870

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT |   |                       |
|--|---|-----------------------|
| Category *   | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
| A  | <p>TEO SOO-HWANG ET AL: "Lifip targets the DNA ligase Lig4p to sites of DNA double-strand breaks."<br/>           CURRENT BIOLOGY,<br/>           vol. 10, no. 3,<br/>           10 February 2000 (2000-02-10), pages<br/>           165-168, XP002178014<br/>           ISSN: 0960-9822<br/>           the whole document</p>      | 1-9                   |
| A  | <p>MILNE G TODD ET AL: "Mutations in two Ku homologs define a DNA end-joining repair pathway in Saccharomyces cerevisiae."<br/>           MOLECULAR AND CELLULAR BIOLOGY,<br/>           vol. 16, no. 8, 1996, pages 4189-4198,<br/>           XP002178015<br/>           ISSN: 0270-7306<br/>           the whole document</p>     | 1-9                   |
| A  | <p>KABOTYANSKI ELENA B ET AL: "Double-strand break repair in Ku86- and XRCC4-deficient cells."<br/>           NUCLEIC ACIDS RESEARCH,<br/>           vol. 26, no. 23,<br/>           1 December 1998 (1998-12-01), pages<br/>           5333-5342, XP002178016<br/>           ISSN: 0305-1048<br/>           the whole document</p> | 1-9                   |
| A,P  | <p>WO 00 68404 A (PIONEER HI BRED INT ;SHI JINRUI (US); MAHAJAN PRAMOD B (US))<br/>           16 November 2000 (2000-11-16)<br/>           the whole document</p>   | 1-9                   |



## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## Continuation of Box I.2

Present claims 1, 25 and 75 relate to an extremely large number of possible methods. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the methods claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the methods, namely: a method to alter a given nucleic acid genomic sequence by inhibiting non-homologous end-joining factors Mre11, Rad50, Nbs1, Lig4, Xrcc4, Ku, and by enhancing the activity of homologous recombination agents as Rad51, Rad52, Rad54 (see claims 2 and 4).

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.



## INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. J. Appl. Application No

PCT/US 01/07870

| Patent document<br>cited in search report |   | Publication<br>date | Patent family<br>member(s)  | Publication<br>date  |
|---|---|---------------------|---|--|
| WO 0012716                                | A | 09-03-2000          | AU 5799599 A<br>EP 1108032 A2<br>WO 0012716 A2<br>US 6180850 B1   | 21-03-2000<br>20-06-2001<br>09-03-2000<br>30-01-2001   |
| WO 9830902                                | A | 16-07-1998          | AU 724108 B2<br>AU 5568198 A<br>AU 729066 B2<br>AU 5568498 A<br>EP 0966683 A1<br>EP 0965040 A1<br>WO 9830902 A1<br>WO 9830903 A1<br>GB 2321702 A ,B<br>GB 2322193 A ,B<br>GB 2329469 A ,B<br>GB 2329248 A ,B<br>JP 2001508868 T<br>JP 2001508175 T<br>US 6242175 B1 | 14-09-2000<br>03-08-1998<br>25-01-2001<br>03-08-1998<br>29-12-1999<br>22-12-1999<br>16-07-1998<br>16-07-1998<br>05-08-1998<br>19-08-1998<br>24-03-1999<br>17-03-1999<br>03-07-2001<br>19-06-2001<br>05-06-2001               |
| WO 9925829                                | A | 27-05-1999          | US 5986055 A<br>AU 1522499 A<br>EP 1030919 A2<br>WO 9925829 A2  | 16-11-1999<br>07-06-1999<br>30-08-2000<br>27-05-1999   |
| WO 9830903                                | A | 16-07-1998          | AU 724108 B2<br>AU 5568198 A<br>AU 729066 B2<br>AU 5568498 A<br>EP 0966683 A1<br>EP 0965040 A1<br>WO 9830902 A1<br>WO 9830903 A1<br>GB 2321702 A ,B<br>GB 2322193 A ,B<br>GB 2329469 A ,B<br>GB 2329248 A ,B<br>JP 2001508868 T<br>JP 2001508175 T<br>US 6242175 B1 | 14-09-2000<br>03-08-1998<br>25-01-2001<br>03-08-1998<br>29-12-1999<br>22-12-1999<br>16-07-1998<br>16-07-1998<br>05-08-1998<br>19-08-1998<br>24-03-1999<br>17-03-1999<br>03-07-2001<br>19-06-2001<br>05-06-2001               |
| US 5641670                                | A | 24-06-1997          | AU 709058 B2<br>AU 2550495 A<br>BR 9507874 A<br>CA 2190289 A1<br>CN 1119545 A<br>CZ 9603258 A3<br>EP 0759082 A1<br>FI 964536 A<br>HU 76844 A2<br>JP 10500570 T<br>NO 964802 A<br>NZ 285945 A<br>SK 146196 A3<br>WO 9531560 A1<br>US 5733746 A<br>ZA 9503879 A       | 19-08-1999<br>05-12-1995<br>19-08-1997<br>23-11-1995<br>03-04-1996<br>17-12-1997<br>26-02-1997<br>09-01-1997<br>28-11-1997<br>20-01-1998<br>09-01-1997<br>25-03-1998<br>04-02-1998<br>23-11-1995<br>31-03-1998<br>18-01-1996 |

Form PCT/ISA/210 (patent family annex) (July 1992)



## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/07870

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date |            |
|---|---------------------|----------------------------|---------------------|------------|
| US 5641670                                | A                   | AU 689455 B2               | 02-04-1998          |            |
|   |                     | AU 5736294 A               | 22-06-1994          |            |
|   |                     | AU 5840198 A               | 04-06-1998          |            |
|   |                     | CA 2151032 A1              | 09-06-1994          |            |
|   |                     | EP 0672160 A1              | 20-09-1995          |            |
|   |                     | JP 8503856 T               | 30-04-1996          |            |
|   |                     | MX 9307623 A1              | 30-06-1994          |            |
|   |                     | NZ 259165 A                | 26-01-1998          |            |
|   |                     | NZ 329348 A                | 28-10-1999          |            |
|   |                     | SG 46476 A1                | 20-02-1998          |            |
|   |                     | TW 402639 B                | 21-08-2000          |            |
|   |                     | WO 9412650 A2              | 09-06-1994          |            |
|   |                     | US 5733761 A               | 31-03-1998          |            |
|   |                     | US 5968502 A               | 19-10-1999          |            |
|   |                     | AT 155810 T                | 15-08-1997          |            |
|   |                     | AU 669499 B2               | 13-06-1996          |            |
|   |                     | AU 3069892 A               | 07-06-1993          |            |
|   |                     | AU 710255 B2               | 16-09-1999          |            |
|   |                     | AU 6563196 A               | 28-11-1996          |            |
|   |                     | DE 69221168 D1             | 04-09-1997          |            |
|   |                     | DE 69221168 T2             | 26-02-1998          |            |
|   |                     | DK 649464 T3               | 02-03-1998          |            |
|   |                     | EP 0649464 A1              | 26-04-1995          |            |
|   |                     | EP 0750044 A2              | 27-12-1996          |            |
|   |                     | ES 2106891 T3              | 16-11-1997          |            |
|   |                     | GR 3025054 T3              | 30-01-1998          |            |
|   |                     | JP 7500969 T               | 02-02-1995          |            |
|   |                     | MX 9206376 A1              | 28-02-1994          |            |
|   |                     | NZ 245015 A                | 21-12-1995          |            |
|   |                     | NZ 270805 A                | 24-10-1997          |            |
| US 5272071                                | A                   | 21-12-1993                 | AT 156189 T         | 15-08-1997 |
|   |                     |                            | AU 645294 B2        | 13-01-1994 |
|   |                     |                            | AU 7183691 A        | 24-07-1991 |
|   |                     |                            | BG 60624 B1         | 31-10-1995 |
|   |                     |                            | BR 9007937 A        | 17-11-1992 |
|   |                     |                            | CA 2071989 A1       | 23-06-1991 |
|   |                     |                            | DE 69031172 D1      | 04-09-1997 |
|   |                     |                            | DE 69031172 T2      | 12-03-1998 |
|   |                     |                            | DE 779362 T1        | 05-04-2001 |
|   |                     |                            | DK 505500 T3        | 25-08-1997 |
|   |                     |                            | EP 0505500 A1       | 30-09-1992 |
|   |                     |                            | EP 0779362 A1       | 18-06-1997 |
|   |                     |                            | ES 2104690 T3       | 16-10-1997 |
|   |                     |                            | ES 2151463 T1       | 01-01-2001 |
|   |                     |                            | FI 922863 A         | 18-06-1992 |
|   |                     |                            | GR 3025057 T3       | 30-01-1998 |
|   |                     |                            | HK 1000547 A1       | 03-04-1998 |
|   |                     |                            | HU 62657 A2         | 28-05-1993 |
|   |                     |                            | HU 217212 B         | 28-12-1999 |
|   |                     |                            | JP 5504682 T        | 22-07-1993 |
|   |                     |                            | KR 176693 B1        | 01-04-1999 |
|   |                     |                            | LT 1595 A ,B        | 25-07-1995 |
|   |                     |                            | LV 10655 A          | 20-04-1995 |
|   |                     |                            | LV 10655 B          | 20-08-1995 |
|   |                     |                            | NO 922436 A         | 19-06-1992 |
|   |                     |                            | OA 9594 A           | 30-04-1993 |
|   |                     |                            | RO 109864 B1        | 30-06-1995 |



## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/07870

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|---|---------------------|----------------------------|---------------------|
| US 5272071 A                              |                     | RU 2128227 C1              | 27-03-1999          |
|   |                     | WO 9109955 A1              | 11-07-1991          |
|   |                     | ZA 9010392 A               | 29-01-1992          |
| WO 0068404 A                              | 16-11-2000          | AU 4975400 A               | 21-11-2000          |
|   |                     | EP 1093523 A1              | 25-04-2001          |
|   |                     | WO 0068404 A1              | 16-11-2000          |



## フロントページの続き

(81)指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

Fターム(参考) 4B024 AA01 AA08 AA10 AA20 BA14  
BA23 BA30 BA80 CA01 CA02  
DA01 DA02 DA03 DA05 DA11  
DA12 FA02 FA06 FA20 GA11  
HA17 HA20  
4B065 AA01X AA58X AA72X AA93X  
AA93Y AB01 AC14 AC20  
BA02 BA16 BA30 CA24 CA43  
CA44 CA46 CA53



## NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

DETAILED DESCRIPTION

---

## [Detailed Description of the Invention]

[0001]

background of invention in of the therapy protein for drugs conveyance (for example, the inside of a vein, hypodermically, or an intramuscular injection) idiomatic on current approach which treats a disease by prescribing therapy protein for the patient vitro production -- and gene therapy is included more recently.

[0002]

The protein aimed at obtaining a therapy can be produced by introducing into a suitable cell the exogenism DNA which carries out the code of the protein aimed at obtaining a therapy. For example, it is possible to introduce into a cell a vector including the exogenism DNA which carries out the code of the therapy protein, and to make the protein by which a code is carried out discover. It has also been suggested by gene targetting that an internal cause sexual cell gene and its manifestation can be embellished. For example, U.S. Pat. No. 5,272,071, U.S. Pat. No. 5,641,670, WO 91/06666, WO 91/06667 and WO Please refer to 90/11354.

Outline of invention Partially, this invention is the part made into a target, approaches a DNA array very much enough, and is based on use of the homologous recombination between the targets DNA DNA, for example, the chromosome in a cell, chosen [ which are chosen and are double-stranded-DNA-arranged ] promoted by offering \*\* 52 which increases homologous recombination, for example, Rad, and \*\* which checks non-homologous edge connection, for example, Ku deactivator. Under existence of both Rad52 and Ku deactivator, it is predicted that homologous recombination of a rate higher than the bottom of nonexistence takes place. Furthermore, gene targetting aiming at changing the part made into the target in DNA, for example, the part made into the target in the chromosome DNA in a cell, using the DNA array chosen as a template is predicted [ that it can promote and ] by offering Rad52 protein and Ku deactivator, for example, anti-Ku antibody. The alteration of a higher rate by gene targetting arises rather than the bottom of the nonexistence of Rad52 protein and Ku deactivator, for example, anti-Ku antibody, by approaching very much the DNA array and target site which are chosen, and offering Rad52 protein and Ku deactivator.

[0003]

Therefore, in one side face, this invention is characterized by the approach of promoting the alteration of the part chosen in Target DNA DNA, for example, the chromosome of a cell. This approach offers \*\* which checks DNA array; and the (c) non-homologous edge connection which carries out the code of \*\* which increases the double-stranded-DNA array;(b) homologous recombination including the DNA array by which :(a) selection is made by the part, for example, Rad52 protein, its functioning fragment, Rad52, or its functioning fragment, for example, \*\* which inactivate Ku, and includes enabling an alteration to happen. The homologous recombination or gene correction between the DNA array as which it sets in a desirable mode, and a part is changed for example, chosen, and Target DNA Since it happens at a rate higher than the rate which will happen under the nonexistence of the homologous recombination improver supplied and a non-homologous edge connection inhibitor By the interaction



part between the DNA array as which the concentration of \*\* which checks the concentration of \*\* and non-homologous edge connection which increase homologous recombination is chosen, and Target DNA, a component (a), (b), and (c) are offered, for example, it is introduced into a cell so that fully. A part is preferably provided with \*\* which checks non-homologous edge connection. Preferably, \*\* which check non-homologous edge connection are Ku deactivators, such as anti-Ku antibody.

[0004]

A component (a), (b), and (c) may be introduced preferably [ both ], or may be introduced separately. Furthermore, both two of components may be introduced and the third component may be introduced separately. For example, both \*\* that may introduce both \*\* 52 that increase a DNA array and homologous recombination, for example, Rad, or check a DNA array and non-homologous edge connection, for example, Ku deactivator, may be introduced. In another desirable mode, both \*\* that check \*\* and non-homologous edge connection which increase homologous recombination may be introduced.

[0005]

two of components -- or all may be preferably offered as complex. \*\* which increases the double-stranded-DNA array;(b) homologous recombination which includes the DNA array by which :(a) selection of this approach is made with Target DNA in a desirable mode, For example, \*\* which checks Rad52 protein or its fragment [ functioning ]; and (c) non-homologous edge connection, For example, it includes contacting the complex containing Ku deactivators, such as anti-Ku antibody, Ku joint oligomer, or a polymer, by introducing for example, this complex into intracellular.

[0006]

In a desirable mode, one or more than it, preferably, components are offered [ no ] by partial conveyance, for example, a microinjection, and are discovered from a target genome or another nucleic acid. It offers by partial conveyance, for example, a microinjection, especially \*\* which checks non-homologous edge connection in a desirable mode, for example, Ku inhibitor, and it is not discovered from a target genome or another nucleic acid.

[0007]

\*\* to which \*\* which checks non-homologous edge connection inactivates :hMre11 in a desirable mode, For example, anti-hMre11 antibody, hMre11 joint oligomer, or a polymer; \*\* which inactivates hRad50, For example, anti-hRad50 antibody, hRad50 joint oligomer, or a polymer; \*\* which inactivates Nbs1, For example, anti-Nbs1 antibody, hNbs1 joint oligomer, or a polymer; \*\* which inactivates the Homo sapiens ligase 4 (hLig4), For example, anti-hLig4 antibody, hLig4 joint oligomer, or a polymer; \*\* which inactivates hXrcc4, For example, anti-hXrcc4 antibody, hXrcc4 joint oligomer, or a polymer; \*\* which inactivates the Homo sapiens homologue (homolog) of Rap1, For example, the oligomer or the polymer combined with the antibody or the Homo sapiens homologue of Rap1 to the Homo sapiens homologue of Rap1; \*\* which inactivates the Homo sapiens homologue of Sir2304, For example, the oligomer or the polymer combined with the antibody or the Homo sapiens homologue of Sir2304 to the Homo sapiens homologue of Sir2304; they are \*\* which inactivates Ku, for example, anti-Ku antibody, Ku joint oligomer, or a polymer. All of \*\* which checks non-homologous edge connection may be prescribed for the patient combining other one or more \*\* which may prescribe a medicine for the patient independently, or check non-homologous edge connection.

[0008]

In a desirable mode, a DNA array is a straight chain DNA array. In a desirable mode, a straight chain DNA array can have one or more single strand overhangs (kind).

[0009]

In a desirable mode, a targeting array adjoins the DNA array chosen. A targeting array is homologous in label, for example, homologous at DNA which adjoins the part which is going to incorporate the part or the DNA array chosen which is going to change Target DNA. Such a contiguity array can be preferably existed [ one or more edges of the DNA array chosen, and ] in both ends. When two contiguity arrays exist, one side should be homologous to the first label field, and another side should be homologous to the second label field.



[0010]

In a desirable mode, a DNA array has one or more protrusion single strand edges, for example, both protrusion both [ one or ] are 3' edge or 5' edges.

In a desirable mode, \*\* which increase homologous recombination are :Rad52 protein, its functioning fragment;Rad51 protein, its functioning fragment;Rad54 protein, functioning fragment;, or those combination.

[0011]

In a desirable mode, \*\* which increases homologous recombination adheres to a DNA array, for example, coating is carried out on a DNA array. In a desirable mode, Rad52 protein or its functioning fragment adheres to a DNA array, for example, coating is carried out on a DNA array.

[0012]

In a desirable mode, Rad52 protein or its fragment is Homo sapiens Rad52 (hRad52).

In a desirable mode, anti-Ku antibody is :anti-Ku70 antibody; anti-Ku80 antibody. In a desirable mode, anti-Ku antibodies are a :hominization antibody; Homo sapiens antibody; antibody fragment, for example, Fab, Fab', and F (ab')<sub>2</sub> or F (v) fragment.

[0013]

In the desirable mode, covalent bond of at least one anti-Ku antibody is carried out to :DNA array;Rad52 protein or its fragment. In another desirable mode, the noncovalent bond of at least one anti-Ku antibody is carried out to :DNA array;Rad52 protein or its fragment.

[0014]

In a desirable mode, anti-Ku70 antibody and anti-Ku80 antibody are offered as a component of complex.

In a desirable mode, a cell is a :eukaryotic cell. In a desirable mode, a cell is the thing of a fungus, vegetation, or the animal origin, for example, the vertebrate origin. The precursor of the cell in which a cell contains :mammalian cell, for example, the founder, or a secondary mammalian cell, for example, fibroblast, a hematopoietic stem cell, myoblast, a keratinization cell, an epithelial cell, an endothelial cell, a neuroglia, a nerve cell, and the element with which blood was formed in a desirable mode, muscular cells, and these somatic cells; they are a transformation or an immortalization cell strain. Preferably, a cell is a human cell. Although not necessarily limited to the example of an immortalization human cell stock useful to this approach, a :Bowes black tumor cell (ATCC deposition number CRL 9607), A Daudi cell (ATCC deposition number CCL 213), a HeLa cell, and a HeLa-cell derivative (derivative) (ATCC deposition number CCL2 CCL2.1 and CCL 2.2), HL-60 cell (ATCC deposition number CCL 240), HT1080 cell (ATCC deposition number CCL 121), A Jurkat cell (ATCC deposition number TIB 152), KB carcinoma cell (ATCC deposition number CCL 17), K-562 leukemic cell (ATCC deposition number CCL 243), MCF-7 cancer cells of breast carcinoma (ATCC deposition number BTH 22), MOLT-4 cell (ATCC deposition number 1582), a Namalwa cell (ATCC deposition number CRL 1432), A Rafji cell (ATCC deposition number CCL 86), RPMI 8226 cells (ATCC deposition number CCL 155), U-937 cell (ATCC deposition number 1593), two RWI-28VA13 low-order stock 4 cell (ATCC deposition number CLL 155), With a CCRF-CEM cell (ATCC deposition number CCL 119) and a 2780AD ovarian cancer cell (Van Der Blick et al., Cancer Res. 48:5927-5932, 1988) The hetero hybridoma cell produced by fusion of a human cell and the cell of another kind is contained. In another mode, immortalization cell strains may be cell strains other than a human cell stock, for example, a CHO cell strain, and a COS cell strain.

[0015]

In a desirable mode, a component, for example, the component of complex, is introduced into a cell by the microinjection.

In one desirable mode, the DNA array chosen differs from Target DNA by 3 or less than 2 [ 10, 8, 6 5, 4, ], one nucleotide, for example, a permutation, deletion, or insertion.

[0016]

In Target DNA, in a desirable mode, a target sequence differs from about 10, 8, 6, 5, 4, 3 and 2 or one nucleotide, and a wild type array, including mutation. Preferably, mutation is point mutation, for



example, insertion, deletion, or the mutation by the permutation.

[0017]

desirable voice -- or [ whether set like, and Target DNA causes this disease or a functional disorder in relation to a disease or a functional disorder, including mutation, or mutation contributes to it, or that it influences it ] -- or it adjusts. A disease or a functional disorder preferably A :cystic-fibrosis; sickle-cell anemia; hemophilia A; hemophilia B; Von Will Brando disease 3 mold; xeroderma pigmentosum; thalassemia; RESSHU-NAIRAN (Lesch-Nylan) syndrome; protein C resistance; lysosome storage disease, For example, Gaucher's disease, Fabry's disease; Mucopolysaccharidosis (MPS) 1 mold (the Harley-SHAIE (Hurley-Scheie) syndrome), MPS II mold (Hunter's syndrome), MPS IIIA mold (Sun Phi RIO (Sanfilio) A syndrome), MPS An IIIB mold (Sun Phi RIO B syndrome), MPS IIIC mold (Sun Phi RIO C syndrome), MPS An IIID mold (Sun Phi RIO D syndrome), MPS IVA mold (MORUKIO A syndrome), MPS An IVB mold (MORUKIO B syndrome), MPS VI mold (MAROTO-rally (Maroteaux-Larry) syndrome), MPS It is a VII mold (Sly's syndrome).

[0018]

In a desirable mode, the DNA array as which Target DNA is chosen, including mutation includes the forward Tsuneno greensand-mold array which can correct mutation.

In a desirable mode, mutation has Target DNA in a cystic-fibrosis film penetration controlling factor (CFTR) gene, including mutation. Preferably, mutation changes the amino acid of the codon 508 of a CFTR protein coding region, and it is 3 base-pair in frame deletion from which mutation removes the phenylalanine of the codon 508 of CFTR protein. This deletion of the phenylalanine -508 in CFTR protein is looked at at a high rate by the test subject who has the cystic fibrosis. Therefore, in a desirable mode, a mutation CFTR gene is made into a target using the DNA array including the array which carries out the code of the phenylalanine -508 which is looked at by the wild type CFTR gene chosen, and correcting is possible.

[0019]

In a desirable mode, mutation has Target DNA in a Homo sapiens beta globin gene, including mutation. Preferably, mutation changes the amino acid of the sixth codon of a beta globin gene, and mutation is the permutation of A [ in / the sixth codon of a beta globin gene ] to T. This mutation draws the change to the valine looked at by the test subject who has sickle-cell anemia from glutamic acid in beta globin protein. Therefore, in a desirable mode, a mutation beta globin gene is made into a target using the DNA array which carries out the code of the wild type amino acid residue by the codon 6, which contains A so that it may see within the sixth codon of DNA chosen, for example, a wild type beta globin gene, and which is chosen, and correcting is possible.

[0020]

In a desirable mode, mutation has Target DNA in a factor VIII gene, including mutation. For example, mutation may be in the exons 23 and 24 of a factor VIII gene, and/or an exon 25. Preferably, mutation changes amino acid by the codon 2209 of the coding region of a factor VIII protein coding region, and it is the permutation of G in the exon 24 of the factor VIII gene to which mutation leads the change to a glutamine from an arginine from the amino acid 2209 of Factor VIII to A. Preferably, mutation changes amino acid by the codon 2229 of the coding region of a factor VIII protein coding region, and it is the permutation of G in the exon 25 of the factor VIII gene to which mutation leads the change to a cysteine from a tryptophan from the amino acid 2229 of Factor VIII to T. Such mutation has been related with serious hemophilia A from whenever [ middle ]. Therefore, in a desirable mode, a mutation factor VIII gene is made into a target using the array containing DNA which carries out the code of the wild type amino acid by the codon 2209 of the coding region of a factor VIII gene, one of DNA which carry out the code of the wild type amino acid by the codon 2229 of the coding region of a factor VIII gene, or both of DNA chosen, and correcting is possible.

[0021]

In a desirable mode, mutation has Target DNA in a factor IX gene, including mutation. For example, in the test subject who has hemophilia B, the great portion of mutation is the point mutation in a factor IX gene. Therefore, in a desirable mode, the DNA array chosen makes a target one or more point mutation



in the factor IX gene relevant to hemophilia B, and in order to correct, one or more nucleotides which have at least one nucleotide of the wild type factor IX gene origin can be included.

[0022]

In a desirable mode, mutation has Target DNA in a von Willebrand factor gene, including mutation. Preferably, mutation is the single cytosine deletion under 6 cytosine expanding of 2679-2684 also in the exon 18 of the Von Will Brando gene. This mutation is looked at at a remarkable rate by the test subject who has Von Will Brando disease 3 mold. Other mutation relevant to Von Will Brando disease 3 mold, for example, point mutation, is changeable so that it may be indicated by this specification. Therefore, in a desirable mode, a mutation Von Will Brando gene is made into a target using the DNA array at least that [ whose ] of the array looked at by the wild type Von Will Brando gene, for example, the Von Will Brando gene, contains six cytosines of 2679-2684 and which is chosen, and correcting is possible.

[0023]

In a desirable mode, mutation has Target DNA in a xeroderma pigmentosum group G (XP-G) gene, including mutation. Preferably, at least that of the 245 base-pair exon as which mutation is regarded by the XP-G gene is the single adenine deletion under 3 adenine expanding of 19-21. This deletion draws a xeroderma pigmentosum. Therefore, in a desirable mode, a mutation XP-G gene is made into a target using DNA in which the wild type array of a XP-G gene, for example, that of a 245 base-pair exon, contains three adenines of 19-21 and which is chosen, and correcting is possible.

[0024]

Preferably, \*\* which inactivates mismatch repair proteins, such as Msh2, Msh6, Msh3, Mlh1, Pms2, Mlh3, and Pms1, is also offered. \*\* may be contained in complex.

[0025]

In another desirable mode, an alteration includes the homologous recombination between the DNA array chosen and Target DNA, for example, a chromosome.

In a desirable mode, the DNA arrays chosen are a sufficient number of nucleotides, and differ from a target so that a target or the DNA array chosen may have a non-pair field, for example, a loop-formation out field, by more nucleotides than 1 unlike Target DNA. In such application, Msh2, Msh6, Msh3, Mlh1, Pms2, Mlh3, and Pms1 may also be offered as some complex.

[0026]

DNA as which an alteration is chosen in a desirable mode, including the nest of the array to Target DNA chosen is incorporated as it has the element on a target chosen beforehand, and the relation chosen beforehand, for example, one side is a control member, and it functions as a control member controlling the manifestation of a protein coding sequence, when it is the array to which another side carries out the code of the protein. The contiguity array which promotes the nest chosen may be used. The DNA array chosen can be incorporated [ 5' of the target sequence chosen, for example, a gene, and a coding sequence, and ] in 3' or this array.

[0027]

In a desirable mode, the DNA array chosen including the nest of the DNA array as which an alteration is chosen is a control array, for example, an exogenous control array. In a desirable mode, a control array includes one or more :promoters and enhancers, an upper activation array (UAS), a scaffold adhesion field, or a transcription factor binding site. In a desirable mode a control array A :metallothionein-I gene, For example, the control array of the mouse metallothionein-I gene origin, the control array of the SV-40 gene origin, The control array of the cytomegalovirus gene origin, the control array of the collagen gene origin, The control array of the actin gene origin, the control array of the immunoglobulin gene origin, The control array of the HMG-CoA reductase gene origin, the control array of gamma actin gene origin, the control array of the transcriptional activator YY1 gene origin, the control array of the fibronectin gene origin, or the control array of the EF-1alpha gene origin is included.

[0028]

In a desirable mode, the DNA array chosen contains an exon. Preferably, an exogenous exon contains the code DNA of the endogenous gene made into :CAP part, a nucleotide sequence ATG, and/or a target, and an in frame.



[0029]

In a desirable mode, the DNA array chosen includes a splice-donor part.

In a desirable mode, when the DNA array chosen is incorporated in label, it includes the exogenous control array which acts so that an endogenous coding sequence may be controlled. The DNA array chosen can be included in the upstream of the coding region of the endogenous gene which can be set in label, or the upstream of the endogenous control array of the endogenous gene which can be set in label. In another desirable mode, the DNA array chosen can be incorporated in an endogenous gene, the lower stream of a river of a coding region, the intron, or an endogenous gene. In another desirable mode, the thing whose endogenous control array of an endogenous gene is inactive and to incorporate is possible for the DNA array chosen so that deletion may be carried out completely partially, for example.

[0030]

In a desirable mode, the DNA array chosen is the upstream of an endogenous gene, and is connected with the second exon of an endogenous gene.

In a desirable mode, an endogenous gene carries out the code of :hormone, cytokine, an antigen, an antibody, an enzyme, a coagulation factor, transporter protein, an acceptor, control protein, structural protein, or the transcription factor. In a desirable mode an endogenous gene The following protein:erythropoietin, Calcitonin, a growth hormone, an insulin, in SURINO tropine, an insulinlike growth factor, Parathyroid hormone, alpha2-interferon (IFNA2), a beta interferon, A gamma interferon, a nerve growth divisor class, FSHbeta, TGF-beta, a tumor necrosis factor, Glucagon, the bone growth factor -2, the bone growth factor -7, TSH-beta, interleukin 1, interleukin 2, interleukin 3, interleukin 6, interleukin 11, interleukin 12, CSF-granulocyte (GCSF), a CSF-macrophage, CSF-granulocyte / macrophage, immunoglobulins, and catalyst nature antibodies Proteinkinase C, a glucocerebrosidase, superoxide dismutase, A tissue plasminogen activator, urokinase, Antithrombin III, DNase, the alpha-galactosidase, tyrosine hydroxylase, a blood coagulation factor V A blood coagulation factor VII, a blood coagulation factor VIII, a blood coagulation factor IX, a blood coagulation factor X A blood coagulation factor XIII, apolipoprotein E, apolipoprotein A-I, Globins, a low-density lipoprotein acceptor, IL-2 acceptor, and IL-2 antagonists An alpha-1-anti trypsin, immunoreaction qualification agents, beta-GURUKO ceramidase, alpha-IZURONIDAZE, alpha-L-IZURONIDAZE, glucosamine-N-sulfatase, Alpha-N-acetyl glucosaminidase, acetyl-coenzyme-A:alpha-glucosamine-N-acetyltransferase, The code of either N-acetyl glucosamine-6-sulfatase, the beta-galactosidase, the beta-glucuronidase, N-acetyl galactosamine-6-sulfatase and the fusibility CD 4 is carried out.

[0031]

In a desirable mode, an endogenous gene carries out the code of the follicle-stimulating hormone beta (FSHbeta), and the DNA array chosen includes the control array from which a control array, for example, the control array and array of an FSHbeta gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of an FSHbeta coding region (array number 1). In a desirable mode, a targeting array is the FSHbeta coding sequence upstream within an FSHbeta coding sequence, and directs homologous recombination. In a desirable mode, a targeting array contains the nucleotide of the array number 3 origin corresponding to -155 which 20, 30, 50,100, or 1000 follow at least from the nucleotide -696 of the nucleotide -7454 of a Homo sapiens FSHbeta array to the array number 2 origin corresponding to -1417 (numbering receives at least a translation open starting area) or a Homo sapiens FSHbeta array.

[0032]

In a desirable mode, an endogenous gene carries out the code of the interferon-alpha 2 (IFNalpha2), and the DNA array chosen includes the control array from which a control array, for example, the control array and array of IFNalpha2 gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of IFNalpha2 coding region. In a desirable mode, a targeting array directs homologous recombination in the field of the IFNalpha2 coding-sequence upstream. In a desirable mode, a targeting array contains the



nucleotide of the array number 4 origin corresponding to -511 (numbering receives at least a translation open starting area) which 20, 30, 50, 100, or 1000 follow at least from the nucleotide -4074 of Homo sapiens IFNalpha2 array. For example This array : From a nucleotide -4074 to at least 20, 30 and 50, or 100 nucleotides of the array number 7 origin corresponding to -3796 of Homo sapiens IFNalpha2 array; At least 20 of the array number 8 origin corresponding to -510 from the nucleotide -582 of Homo sapiens IFNalpha2 array, 30, Or 50 nucleotides; at least 20, 30, 50 and 100, or 1000 nucleotides of the array number 9 origin corresponding to -583 may also be included from the nucleotide -3795 of Homo sapiens IFNalpha2 array.

[0033]

In a desirable mode, an endogenous gene carries out the code of the granulocyte colony-stimulating factor (GCSF), and the DNA array chosen includes the control array from which a control array, for example, the control array and array of a GCSF gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of a GCSF coding region. In a desirable mode, a targeting array is the GCSF coding sequence upstream within a GCSF coding sequence, and directs homologous recombination. In a desirable mode, a targeting array contains the nucleotide of the array number 5 origin corresponding to 101 (numbering receives at least a translation open starting area) which 20, 30, 50, 100, or 1000 follow at least from the nucleotide -6,578 of a Homo sapiens GCSF array. For example, a target sequence may also contain 20, 30, 50, 100, or 1000 nucleotides of the array number 6 origin corresponding to -364 from the nucleotide -6,578 of a Homo sapiens GCSF gene.

[0034]

In another desirable mode, the DNA array as which a DNA array is chosen, including a coding region carries out the code of the protein. In a desirable mode, a coding region carries out the code of :hormone, cytokine, an antigen, an antibody, an enzyme, a coagulation factor, transporter protein, an acceptor, control protein, structural protein, or the transcription factor. In a desirable mode a coding region The following protein:erythropoietin, Calcitonin, a growth hormone, an insulin, in SURINO tropine, an insulinlike growth factor, Parathyroid hormone, alpha2-interferon (IFNA2), a beta interferon, A gamma interferon, a nerve growth divisor class, FSHbeta, TGF-beta, a tumor necrosis factor, Glucagon, the bone growth factor -2, the bone growth factor -7, TSH-beta, interleukin 1, interleukin 2, interleukin 3, interleukin 6, interleukin 11, interleukin 12, CSF-granulocyte (GCSF), a CSF-macrophage, CSF-granulocyte / macrophage, immunoglobulins, and catalyst nature antibodies Proteinkinase C, a glucocerebrosidase, superoxide dismutase, A tissue plasminogen activator, urokinase, Antithrombin III, DNase, the alpha-galactosidase, tyrosine hydroxylase, a blood coagulation factor V A blood coagulation factor VII, a blood coagulation factor VIII, a blood coagulation factor IX, a blood coagulation factor X A blood coagulation factor XIII, apolipoprotein E, apolipoprotein A-I, Globins, a low-density lipoprotein acceptor, IL-2 acceptor, and IL-2 antagonists An alpha-1-anti trypsin, immunoreaction qualification agents, beta-GURUKO ceramidase, alpha-IZURONIDAZE, alpha-L-IZURONIDAZE, glucosamine-N-sulfatase, Alpha-N-acetyl glucosaminidase, acetyl-coenzyme-A:alpha-glucosamine-N-acetyltransferase, The code of either N-acetyl glucosamine-6-sulfatase, the beta-galactosidase, the beta-glucuronidase, N-acetyl galactosamine-6-sulfatase and the fusibility CD 4 is carried out.

[0035]

In a desirable mode, as the DNA array chosen is under accommodation of an endogenous control member, it is possible to incorporate this array in a target. DNA chosen can be included in the lower stream of a river of an endogenous control array or the upstream of the coding region of an endogenous gene, and the lower stream of a river of the endogenous control array of a gene. In another desirable mode, DNA chosen can be included in the lower stream of a river of an endogenous control array so that deletion of the coding region of an endogenous gene may be inactivated for example, carried out completely partially.

[0036]

desirable voice -- like -- setting -- this approach -- a mismatch repair protein, Msh2, Msh6, and Msh3,



M1h1, and Pms2 and M1 -- it includes further introducing \*\* which checks h3, Pms1, other mismatch repair proteins, or those combination. [ for example, ] Preferably, \*\* is \*\* which checks the manifestation of a mismatch repair protein, for example, \*\* is an antisense RNA. In a desirable mode, \*\* is an antibody to a mismatch repair protein. desirable voice -- an antibody [ as opposed to / set like and / a mismatch repair protein ] -- complex -- a share -- or the noncovalent bond is carried out.

[0037]

In another side face, this invention is characterized by the constituent for promoting the alteration by Target's DNA, for example, a chromosome, for example, this specification publication, target DNA, for example, the complex of a component, using as a template the DNA array chosen, for example, the DNA array as which this specification publication is chosen. A constituent contains \*\* which checks \*\* which increases the double-stranded-DNA array;(b) homologous recombination including DNA array by which :(a) selection is made, for example, Rad52 protein, and its fragment [ functioning ];, and (c) non-homologous edge connection, for example, \*\* which inactivates Ku. A target DNA array is changeable with a nest, using this constituent.

[0038]

\*\* to which \*\* which checks non-homologous edge connection inactivates :hMre11 in a desirable mode, For example, anti-hMre11 antibody, hMre11 joint oligomer, or a polymer; \*\* which inactivates hRad50, For example, anti-hRad50 antibody, hRad50 joint oligomer, or a polymer; \*\* which inactivates Nbs1, For example, anti-Nbs1 antibody, hNbs1 joint oligomer, or a polymer; \*\* which inactivates the Homo sapiens ligase 4 (hLig4), For example, anti-hLig4 antibody, hLig4 joint oligomer, or a polymer; \*\* which inactivates hXrcc4, For example, anti-hXrcc4 antibody, hXrcc4 joint oligomer, or a polymer; \*\* which inactivates the Homo sapiens homologue of Rap1, For example, the oligomer or the polymer combined with the antibody or the Homo sapiens homologue of Rap1 to the Homo sapiens homologue of Rap1; \*\* which inactivates the Homo sapiens homologue of Sir2304, For example, the oligomer or the polymer combined with the antibody or the Homo sapiens homologue of Sir2304 to the Homo sapiens homologue of Sir2304; they are \*\* which inactivates Ku, for example, anti-Ku antibody, Ku joint oligomer, or a polymer. All of \*\* which checks non-homologous edge connection may be prescribed for the patient combining other one or more \*\* which may prescribe a medicine for the patient independently, or check non-homologous edge connection.

[0039]

In a desirable mode, a DNA array is a straight chain DNA array. In a desirable mode, a straight chain DNA array can have one or more single strand overhangs (kind).

[0040]

In a desirable mode, a targeting array adjoins the DNA array chosen. A targeting array is homologous in label, for example, homologous at DNA which adjoins the part which is going to insert the part or the DNA array chosen which is going to change Target DNA. Such a contiguity array can be preferably existed [ one or more edges of the DNA array chosen, and ] in both ends. When two contiguity arrays exist, one side should be homologous to the first label field, and another side should be homologous to the second label field.

[0041]

In a desirable mode, a DNA array has one or more protrusion single strand edges, for example, both protrusion both [ one or ] are 3' edge or 5' edges.

In a desirable mode, \*\* which increase homologous recombination are :Rad52 protein, its functioning fragment;Rad51 protein, its functioning fragment;Rad54 protein, functioning fragment;; or those combination.

[0042]

In a desirable mode, \*\* which increases homologous recombination adheres to a DNA array, for example, coating is carried out on a DNA array. In a desirable mode, coating of Rad52 protein or its functioning fragment is carried out on the DNA array which adheres to the DNA array chosen, for example, is chosen.

[0043]



In a desirable mode, Rad52 protein or its fragment is Homo sapiens Rad52 (hRad52).

In a desirable mode, anti-Ku antibody is :anti-Ku70 antibody; anti-Ku80 antibody. In a desirable mode, anti-Ku antibodies are a :hominization antibody; Homo sapiens antibody; antibody fragment, for example, Fab, Fab', and F (ab')<sub>2</sub> or F (v) fragment.

[0044]

In the desirable mode, covalent bond of at least one anti-Ku antibody is carried out to the DNA array;Rad52 protein of which :selection is done, or its fragment. In another desirable mode, the noncovalent bond of at least one anti-Ku antibody is carried out to the DNA array;Rad52 protein of which :selection is done, or its fragment.

[0045]

In a desirable mode, a constituent contains anti-Ku70 antibody and anti-Ku80 antibody.

In a desirable mode, the DNA array chosen differs from Target DNA by 3 or less than 2 [ 10, 8, 6 5, 4, ], one nucleotide, for example, a permutation, deletion, or insertion.

[0046]

In Target DNA, in a desirable mode, a target sequence differs from about 10, 8, 6, 5, 4, 3 and 2 or one nucleotide, and a wild type array, including mutation. Preferably, mutation is point mutation, for example, insertion, deletion, or the mutation by the permutation.

[0047]

desirable voice -- or [ whether set like, and Target DNA causes this disease or a functional disorder in relation to a disease or a functional disorder, including mutation, or mutation contributes to it, or that it influences it ] -- or it adjusts. A disease or a functional disorder preferably A :cystic-fibrosis; sickle-cell anemia; hemophilia A; hemophilia B; Von Will Brando disease 3 mold; xeroderma pigmentosum; thalassemia; RESSHU-NAIRAN syndrome; protein C resistance; lysosome storage disease, for example, Gaucher's disease, Fabry's disease; Mucopolysaccharidosis (MPS) 1 mold (Harley-Scheie's syndrome), MPS II mold (Hunter's syndrome), MPS IIIA mold (Sun Phi RIO A syndrome), MPS An IIIB mold (Sun Phi RIO B syndrome), MPS IIIC mold (Sun Phi RIO C syndrome), MPS An IIID mold (Sun Phi RIO D syndrome), MPS An IVA mold (MORUKIO A syndrome), MPS An IVB mold (MORUKIO B syndrome), MPS VI mold (MAROTO-rally syndrome), MPS It is a VII mold (Sly's syndrome).

[0048]

In a desirable mode, the DNA array as which Target DNA is chosen, including mutation includes the forward Tsuneno greensand-mold array which can correct mutation.

In a desirable mode, mutation has Target DNA in a cystic-fibrosis film penetration controlling factor (CFTR) gene, including mutation. Preferably, mutation changes the amino acid of the codon 508 of a CFTR protein coding region, and it is 3 base-pair in frame deletion from which mutation removes the phenylalanine of the codon 508 of CFTR protein. This deletion of the phenylalanine -508 in CFTR protein is looked at at a high rate by the test subject who has the cystic fibrosis. Therefore, in a desirable mode, a mutation CFTR gene is made into a target using the DNA array including the array which carries out the code of the phenylalanine -508 which is looked at by the wild type CFTR gene chosen, and correcting is possible.

[0049]

In a desirable mode, mutation has Target DNA in a Homo sapiens beta globin gene, including mutation. Preferably, mutation changes the amino acid of the sixth codon of a beta globin gene, and mutation is the permutation of A [ in / the sixth codon of a beta globin gene ] to T. This mutation draws the change to the valine looked at by the test subject who has sickle-cell anemia from glutamic acid in beta globin protein. Therefore, in a desirable mode, a mutation beta globin gene is made into a target using the DNA array which carries out the code of the wild type amino acid residue by the codon 6, which contains A so that it may see within the sixth codon of DNA chosen, for example, a wild type beta globin gene, and which is chosen, and correcting is possible.

[0050]

In a desirable mode, mutation has Target DNA in a factor VIII gene, including mutation. For example,



mutation may be in the exons 23 and 24 of a factor VIII gene, and/or an exon 25. Preferably, mutation changes amino acid by the codon 2209 of the coding region of a factor VIII protein coding region, and it is the permutation of G in the exon 24 of the factor VIII gene to which mutation leads the change to a glutamine from an arginine from the amino acid 2209 of Factor VIII to A. Preferably, mutation changes amino acid by the codon 2229 of the coding region of a factor VIII protein coding region, and it is the permutation of G in the exon 25 of the factor VIII gene to which mutation leads the change to a cysteine from a tryptophan from the amino acid 2229 of Factor VIII to T. Such mutation has been related with serious hemophilia A from whenever [ middle ]. Therefore, in a desirable mode, a mutation factor VIII gene is made into a target using the DNA array containing DNA which carries out the code of the wild type amino acid by the codon 2209 of the coding region of a factor VIII gene, one of DNA which carry out the code of the wild type amino acid by the codon 2229 of the coding region of a factor VIII gene, or both chosen, and correcting is possible.

[0051]

In a desirable mode, mutation has Target DNA in a factor IX gene, including mutation. For example, in the test subject who has hemophilia B, the great portion of mutation is the point mutation in a factor IX gene. Therefore, in a desirable mode, the DNA array chosen makes a target one or more point mutation in the factor IX gene relevant to hemophilia B, and in order to correct, one or more nucleotides which have at least one nucleotide of the wild type factor IX gene origin can be included.

[0052]

In a desirable mode, mutation has Target DNA in a von Willebrand factor gene, including mutation. Preferably, mutation is the single cytosine deletion under 6 cytosine expanding of 2679-2684 also in the exon 18 of the Von Will Brando gene. This mutation is looked at at a remarkable rate by the test subject who has Von Will Brando disease 3 mold. Other mutation relevant to Von Will Brando disease 3 mold, for example, point mutation, is changeable so that it may be indicated by this specification. Therefore, in a desirable mode, a mutation Von Will Brando gene is made into a target using the DNA array at least that [ whose ] of the exon 18 of the array looked at by the wild type Von Will Brando gene, for example, the Von Will Brando gene, contains six cytosines of 2679-2684 and which is chosen, and correcting is possible.

[0053]

In a desirable mode, mutation has Target DNA in a xeroderma pigmentosum group G (XP-G) gene, including mutation. Preferably, at least that of the 245 base-pair exon as which mutation is regarded by the XP-G gene is the single adenine deletion under 3 adenine expanding of 19-21. This deletion draws a xeroderma pigmentosum. Therefore, in a desirable mode, a mutation XP-G gene is made into a target using DNA in which the wild type array of a XP-G gene, for example, that of the 245 base-pair exon of a XP-G gene, contains three adenines of 19-21 and which is chosen, and correcting is possible.

[0054]

In another desirable mode, the DNA arrays chosen are a sufficient number of nucleotides, and differ from a target so that a target or the DNA array chosen may have a non-pair field, for example, a loop-formation out field, by more nucleotides than 1 unlike Target DNA. Preferably, \*\* which inactivates mismatch repair proteins, such as Msh2, Msh6, Msh3, Mlh1, Pms2, Mlh3, Pms(es)1, or those combination, may also be included in a constituent, for example, may also contain this \*\* in complex.

[0055]

In a desirable mode, the DNA array chosen has a contiguity array so that it may be possible the element on Target DNA chosen beforehand and to be incorporated as it has the relation chosen beforehand. For example, DNA chosen is a control array, and the contiguity array will be incorporated so that it may function as a control member controlling the manifestation of a protein coding sequence, when Target DNA does the code of the protein. It is possible to use the contiguity array which promotes the nest chosen. The DNA array chosen can have the gene of the target sequence chosen, for example, label hit, or 5' of a coding region, and the contiguity array that can be incorporated in 3' or this field.

[0056]

In a desirable mode, the DNA array chosen includes a control array, for example, an exogenous control



array. In a desirable mode, a control array includes one or more promoters and enhancers, UAS, a scaffold adhesion field, or a transcription factor binding site. In a desirable mode a control array A :metallothionein-I gene, For example, the control array of the mouse metallothionein-I gene origin, the control array of the SV-40 gene origin, The control array of the cytomegalovirus gene origin, the control array of the collagen gene origin, The control array of the actin gene origin, the control array of the immunoglobulin gene origin, The control array of the HMG-CoA reductase gene origin, the control array of gamma actin gene origin, the control array of the transcriptional activator YY1 gene origin, the control array of the fibronectin gene origin, or the control array of the EF-1alpha gene origin is included.

[0057]

In a desirable mode, the DNA array chosen contains an exon. Preferably, an exogenous exon contains the code DNA of the endogenous gene made into :CAP part, a nucleotide sequence ATG, and/or a target, and an in frame.

[0058]

In a desirable mode, the DNA array chosen includes a splice-donor part.

In a desirable mode, a constituent including the DNA array which has an exogenous control array and which is chosen may have a contiguity array which is incorporated in a target so that it may function as the DNA array chosen controlling the manifestation of an endogenous array. It sets in label and DNA chosen can be incorporated like a label of an endogenous gene or the coding region upstream of a coding sequence, or it sets in label and it can be incorporated like a label of an endogenous gene or the endogenous control array upstream of a coding sequence. In another desirable mode, the thing whose endogenous control array of an endogenous gene is inactive and to include in label hit is possible for the DNA array chosen so that deletion may be carried out completely partially, for example. The DNA array chosen can be incorporated like a label of the lower stream of a river of an endogenous gene or a coding region, or can be incorporated in the intron of an endogenous gene.

[0059]

In a desirable mode, the DNA array chosen includes the control array from which a control array, for example, the control array and array of an FSHbeta gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of an FSHbeta coding region (array number 1). In a desirable mode, a targeting array is the FSHbeta coding sequence upstream within an FSHbeta coding sequence, and directs homologous recombination. In a desirable mode, a targeting array contains the nucleotide of the array number 3 origin corresponding to -155 which 20, 30, 50,100, or 1000 follow at least from the nucleotide -696 of the nucleotide -7454 of a Homo sapiens FSHbeta array to the array number 2 origin corresponding to -1417 (numbering receives at least a translation open starting area) or a Homo sapiens FSHbeta array.

[0060]

In a desirable mode, the DNA array chosen includes the control array from which a control array, for example, the control array and array of IFNalpha2 gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of IFNalpha2 coding region. In a desirable mode, a targeting array directs homologous recombination in the field of the IFNalpha2 coding-sequence upstream. In a desirable mode, a targeting array contains the nucleotide of the array number 4 origin corresponding to -511 (numbering receives at least a translation open starting area) which 20, 30, 50,100, or 1000 follow at least from the nucleotide -4074 of Homo sapiens IFNalpha2 array. For example This array : From a nucleotide -4074 to at least 20, 30 and 50, or 100 nucleotides of the array number 7 origin corresponding to -3796 of Homo sapiens IFNalpha2 array; At least 20 of the array number 8 origin corresponding to -510 from the nucleotide -582 of Homo sapiens IFNalpha2 array, 30, Or 50 nucleotides; at least 20, 30, 50 and 100, or 1000 nucleotides of the array number 9 origin corresponding to -583 may also be included from the nucleotide -3795 of Homo sapiens IFNalpha2 array.



[0061]

In a desirable mode, the DNA array chosen includes the control array from which a control array, for example, the control array and array of a GCSF gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of a GCSF coding region. In a desirable mode, a targeting array is; or the GCSF coding sequence upstream within :GCSF coding sequence, and directs homologous recombination. In a desirable mode, a targeting array contains the nucleotide of the array number 5 origin corresponding to 101 (numbering receives at least a translation open starting area) which 20, 30, 50,100, or 1000 follow at least from the nucleotide -6,578 of a Homo sapiens GCSF array. For example, a target sequence may also contain 20, 30, 50,100, or 1000 nucleotides of the array number 6 origin corresponding to -364 (numbering receives at least a translation open starting area) from the nucleotide -6,578 of a Homo sapiens GCSF gene.

[0062]

In another desirable mode, as for a DNA array, a DNA array carries out the code of the protein, including a coding region. In a desirable mode, a coding region carries out the code of :hormone, cytokine, an antigen, an antibody, an enzyme, a coagulation factor, transporter protein, an acceptor, control protein, structural protein, or the transcription factor. In a desirable mode a coding region The following protein:erythropoietin, Calcitonin, a growth hormone, an insulin, in SURINO tropine, an insulinlike growth factor, Parathyroid hormone, alpha2-interferon (IFNA2), a beta interferon, A gamma interferon, a nerve growth divisor class, FSHbeta, TGF-beta, a tumor necrosis factor, Glucagon, the bone growth factor -2, the bone growth factor -7, TSH-beta, interleukin 1, interleukin 2, interleukin 3, interleukin 6, interleukin 11, interleukin 12, CSF-granulocyte (GCSF), a CSF-macrophage, CSF-granulocyte / macrophage, immunoglobulins, and catalyst nature antibodies Proteinkinase C, a glucocerebrosidase, superoxide dismutase, A tissue plasminogen activator, urokinase, Antithrombin III, DNase, the alpha-galactosidase, tyrosine hydroxylase, a blood coagulation factor V A blood coagulation factor VII, a blood coagulation factor VIII, a blood coagulation factor IX, a blood coagulation factor X A blood coagulation factor XIII, apolipoprotein E, apolipoprotein A-I, Globins, a low-density lipoprotein acceptor, IL-2 acceptor, and IL-2 antagonists An alpha-1-anti trypsin, immunoreaction qualification agents, beta-GURUKO ceramidase, alpha-IZURONIDAZE, alpha-L-IZURONIDAZE, glucosamine-N-sulfatase, Alpha-N-acetyl glucosaminidase, acetyl-coenzyme-A:alpha-glucosamine-N-acetyltransferase, The code of either N-acetyl glucosamine-6-sulfatase, the beta-galactosidase, the beta-glucuronidase, N-acetyl galactosamine-6-sulfatase and the fusibility CD 4 is carried out.

[0063]

In a desirable mode, in case the DNA array chosen is incorporated in a target, as it is under accommodation of an endogenous control member, you may have a contiguity array. DNA chosen can be included in the lower stream of a river of an endogenous control array or the upstream of the coding region of an endogenous gene, and the lower stream of a river of the endogenous control array of a gene. In another desirable mode, DNA chosen can be included in the lower stream of a river of an endogenous control array so that deletion of the coding region of an endogenous gene may be inactivated for example, carried out completely partially.

[0064]

A constituent, for example, complex, is introduced into a cell in a desirable mode. Preferably, a cell is an eukaryotic cell. In a desirable mode, a cell is the thing of a fungus, vegetation, or the animal origin, for example, the vertebrate origin. The precursor of the cell in which a cell contains :mammalian cell, for example, the founder, or a secondary mammalian cell, for example, fibroblast, a hematopoietic stem cell, myoblast, a keratinization cell, an epithelial cell, an endothelial cell, a neuroglia, a nerve cell, and the element with which blood was formed in a desirable mode, muscular cells, and these somatic cells; they are a transformation or an immortalization cell strain. Preferably, a cell is a human cell. Although not necessarily limited to the example of an immortalization human cell stock useful to this approach, a :Bowes black tumor cell (ATCC deposition number CRL 9607), A Daudi cell (ATCC deposition number CCL 213), a HeLa cell, and a HeLa-cell derivative (ATCC deposition number CCL2 CCL2.1



and CCL 2.2), HL-60 cell (ATCC deposition number CCL 240), HT1080 cell (ATCC deposition number CCL 121), A Jurkat cell (ATCC deposition number TIB 152), KB carcinoma cell (ATCC deposition number CCL 17), K-562 leukemic cell (ATCC deposition number CCL 243), MCF-7 cancer cells of breast carcinoma (ATCC deposition number BTH 22), MOLT-4 cell (ATCC deposition number 1582), a Namalwa cell (ATCC deposition number CRL 1432), A Raji cell (ATCC deposition number CCL 86), RPMI 8226 cells (ATCC deposition number CCL 155), U-937 cell (ATCC deposition number 1593), two RWI-28VA13 low-order stock 4 cell (ATCC deposition number CLL 155), With a CCRF-CEM cell (ATCC deposition number CCL 119) and a 2780AD ovarian cancer cell (Van Der Blick et al., Cancer Res. 48:5927-5932, 1988) The hetero hybridoma cell produced by fusion of a human cell and the cell of another kind is contained. In another mode, immortalization cell strains may be cell strains other than a human cell stock, for example, a CHO cell strain, and a COS cell strain.

[0065]

desirable voice -- like -- setting -- a constituent -- a mismatch repair protein, Msh2, Msh6, and Msh3, Mlh1, and Pms2 and M1 -- \*\* which checks h3, Pms1, other mismatch repair proteins, or those combination is included further. [ for example, ] Preferably, \*\* is \*\* which checks the manifestation of a mismatch repair protein, for example, \*\* is an antisense RNA. In a desirable mode, \*\* is an antibody to a mismatch repair protein. desirable voice -- an antibody [ as opposed to / set like and / a mismatch repair protein ] -- one or more components of a constituent -- a share -- or the noncovalent bond is carried out.

[0066]

In another side face, this invention is characterized by the approach of offering protein. This approach offers the cell created by the approach indicated by :book specification, and it includes enabling a cell to discover protein.

[0067]

In another desirable mode, the :this approach to the part made into a target for :alteration \*\* which increases the double-stranded-DNA array;(b) homologous recombination including the DNA array by which the following component:(a) selections are made, For example, the cell into which \*\* which checks Rad52 protein or its fragment [ functioning ]; and (c) non-homologous edge connection, for example, Ku deactivator, is introduced is offered, and it includes enabling a cell to discover protein. A proteinic manifestation can be happened enabling the proteinic manifestation by which a code is carried out to DNA, or by activating a proteinic manifestation.

[0068]

The homologous recombination or gene correction between the DNA array as which it sets in a desirable mode, and a part is changed for example, chosen, and Target DNA Since it happens at a rate higher than the rate which will happen under the nonexistence of the homologous recombination improver supplied and a non-homologous edge connection inhibitor By the interaction part between the DNA array as which the concentration of \*\* which checks the concentration of \*\* and non-homologous edge connection which increase homologous recombination is chosen, and Target DNA, a component (a), (b), and (c) are offered, for example, it is introduced into a cell so that fully. A part is preferably provided with \*\* which checks non-homologous edge connection.

[0069]

In a desirable mode, both a component (a), (b) and, and (c) may be introduced, or may be introduced separately. Furthermore, both two of components may be introduced and the third component may be introduced separately. For example, both \*\* that may introduce both \*\* 52 that increase a DNA array and homologous recombination, for example, Rad, or check a DNA array and non-homologous edge connection, for example, Ku deactivator, may be introduced. In another desirable mode, both \*\* that check \*\* and non-homologous edge connection which increase homologous recombination may be introduced.

[0070]

two of components -- or all may be preferably offered as complex. It includes contacting the complex containing \*\* which checks \*\* which increases the double-stranded-DNA array;(b) homologous



recombination including DNA array by which : (a) selection of this approach is made with Target DNA in desirable mode, for example, Rad52 protein, and its fragment [ functioning ], and (c) non-homologous edge connection, for example, \*\* which inactivates Ku, by introducing for example, this complex into intracellular.

[0071]

In a desirable mode, one or more than it, preferably, components are offered [ no ] by partial conveyance, for example, a microinjection, and are discovered from a target genome or other nucleic acids. Especially Ku deactivators, such as \*\* which checks non-homologous edge connection in a desirable mode, for example, anti-Ku antibody etc., are offered by partial conveyance, for example, a microinjection, and are not discovered from a target genome or other nucleic acids.

[0072]

\*\* to which \*\* which checks non-homologous edge connection inactivates :hMre11 in a desirable mode, For example, anti-hMre11 antibody, hMre11 joint oligomer, or a polymer; \*\* which inactivates hRad50, For example, anti-hRad50 antibody, hRad50 joint oligomer, or a polymer; \*\* which inactivates Nbs1, For example, anti-Nbs1 antibody, hNbs1 joint oligomer, or a polymer; \*\* which inactivates the Homo sapiens ligase 4 (hLig4), For example, anti-hLig4 antibody, hLig4 joint oligomer, or a polymer; \*\* which inactivates hXrcc4, For example, anti-hXrcc4 antibody, hXrcc4 joint oligomer, or a polymer; \*\* which inactivates the Homo sapiens homologue of Rap1, For example, the oligomer or the polymer combined with the antibody or the Homo sapiens homologue of Rap1 to the Homo sapiens homologue of Rap1; \*\* which inactivates the Homo sapiens homologue of Sir2304, For example, the oligomer or the polymer combined with the antibody or the Homo sapiens homologue of Sir2304 to the Homo sapiens homologue of Sir2304; they are \*\* which inactivates Ku, for example, anti-Ku antibody, Ku joint oligomer, or a polymer. All of \*\* which checks non-homologous edge connection may be prescribed for the patient combining other one or more \*\* which may prescribe a medicine for the patient independently, or check non-homologous edge connection.

[0073]

In a desirable mode, a DNA array is a straight chain DNA array. In a desirable mode, a straight chain DNA array can have one or more single strand overhangs (kind).

[0074]

In a desirable mode, a targeting array adjoins the DNA array chosen. A targeting array is homologous in label, for example, homologous at DNA which adjoins the part which is going to incorporate the part or the DNA array chosen which is going to change Target DNA. Such a contiguity array can be preferably existed [ one or more edges of the DNA array chosen, and ] in both ends. When two contiguity arrays exist, one side should be homologous to the first label field, and another side should be homologous to the second label field.

[0075]

In a desirable mode, a DNA array has one or more protrusion single strand edges, for example, both protrusion both [ one or ] are 3' edge or 5' edges.

In a desirable mode, \*\* which increase homologous recombination are :Rad52 protein, its functioning fragment;Rad51 protein, its functioning fragment;Rad54 protein, functioning fragment;, or those combination.

[0076]

In a desirable mode, \*\* which increases homologous recombination adheres to a DNA array, for example, coating is carried out on a DNA array. In a desirable mode, coating of Rad52 protein or its functioning fragment is carried out on the DNA array which adheres to the DNA array chosen, for example, is chosen.

[0077]

In a desirable mode, Rad52 protein or its fragment is Homo sapiens Rad52 (hRad52).

In a desirable mode, anti-Ku antibody is :anti-Ku70 antibody; anti-Ku80 antibody. In a desirable mode, anti-Ku antibodies are a :hominization antibody; Homo sapiens antibody; antibody fragment, for example, Fab, Fab', and F (ab')<sub>2</sub> or F (v) fragment.



[0078]

The DNA array by which :selection of at least one anti-Ku antibody is done in a desirable mode; covalent bond is carried out to \*\* which increases homologous recombination, for example, Rad52 protein, and its fragment. The DNA array by which :selection of at least one anti-Ku antibody is done in another desirable mode; the noncovalent bond is carried out to \*\* which increases homologous recombination, for example, Rad52 protein, and its fragment.

[0079]

In a desirable mode, complex contains anti-Ku70 antibody and anti-Ku80 antibody which are offered as a component of complex.

In a desirable mode, a cell is a :eukaryotic cell. In a desirable mode, a cell is the thing of a fungus, vegetation, or the animal origin, for example, the vertebrate origin. The precursor of the cell in which a cell contains :mammalian cell, for example, the founder, or a secondary mammalian cell, for example, fibroblast, a hematopoietic stem cell, myoblast, a keratinization cell, an epithelial cell, an endothelial cell, a neuroglia, a nerve cell, and the element with which blood was formed in a desirable mode, muscular cells, and these somatic cells; they are a transformation or an immortalization cell strain. Preferably, a cell is a human cell. Although not necessarily limited to the example of an immortalization human cell stock useful to this approach, a :Bowes black tumor cell (ATCC deposition number CRL 9607), A Daudi cell (ATCC deposition number CCL 213), a HeLa cell, and a HeLa-cell derivative (ATCC deposition number CCL2 CCL2.1 and CCL 2.2), HL-60 cell (ATCC deposition number CCL 240), HT1080 cell (ATCC deposition number CCL 121), A Jurkat cell (ATCC deposition number TIB 152), KB carcinoma cell (ATCC deposition number CCL 17), K-562 leukemic cell (ATCC deposition number CCL 243), MCF-7 cancer cells of breast carcinoma (ATCC deposition number BTH 22), MOLT-4 cell (ATCC deposition number 1582), a Namalwa cell (ATCC deposition number CRL 1432), A Rafji cell (ATCC deposition number CCL 86), RPMI 8226 cells (ATCC deposition number CCL 155), U-937 cell (ATCC deposition number 1593), two RWI-28VA13 low-order stock 4 cell (ATCC deposition number CLL 155), With a CCRF-CEM cell (ATCC deposition number CCL 119) and a 2780AD ovarian cancer cell (Van Der Blick et al., Cancer Res. 48:5927-5932, 1988) The hetero hybridoma cell produced by fusion of a human cell and the cell of another kind is contained. In another mode, immortalization cell strains may be cell strains other than a human cell stock, for example, a CHO cell strain, and a COS cell strain.

[0080]

In a desirable mode, a component, for example, the component of complex, is introduced into a cell by the microinjection.

desirable voice -- like -- setting -- an approach -- a mismatch repair protein, Msh2, Msh6, and Msh3, Mlh1, and Pms2 and M1 -- it includes further introducing \*\* which checks h3, Pms1, other mismatch repair proteins, or those combination. [ for example, ] Preferably, \*\* is \*\* which checks the manifestation of a mismatch repair protein, for example, \*\* is an antisense RNA. In a desirable mode, \*\* is an antibody to a mismatch repair protein. desirable voice -- an antibody [ as opposed to / set like and / a mismatch repair protein ] -- complex -- a share -- or the noncovalent bond is carried out.

[0081]

Setting in a desirable mode, protein is in. It is discovered by vitro. In other desirable modes, a cell is offered in a test subject, for example, Homo sapiens, and protein is discovered in a test subject. In a desirable mode, protein is discovered in a test subject and cells are self, allogeneic, and different species. DNA chosen is in. It is ex about DNA which may introduce into a cell by vivo, or removes a cell from a test subject and is chosen. It may introduce by vivo and a cell may be returned to a test subject.

[0082]

In a desirable mode, the DNA array chosen differs from Target DNA by 3 or less than 2 [ 10, 8, 6 5, 4, ], one nucleotide, for example, a permutation, deletion, or insertion.

[0083]

In Target DNA, in a desirable mode, a target sequence differs from about 10, 8, 6, 5, 4, 3 and 2 or one nucleotide, and a wild type array, including mutation. Preferably, mutation is point mutation, for



example, insertion, deletion, or the mutation by the permutation.

[0084]

desirable voice -- or [ whether set like, and Target DNA causes this disease or a functional disorder in relation to a disease or a functional disorder, including mutation, or mutation contributes to it, or that it influences it ] -- or it adjusts. A disease or a functional disorder preferably A :cystic-fibrosis; sickle-cell anemia; hemophilia A; hemophilia B; Von Will Brando disease 3 mold; xeroderma pigmentosum; thalassemia; RESSHU-NAIRAN syndrome; protein C resistance; lysosome disease, For example, Gaucher's disease, Fabry's disease; Mucopolysaccharidosis (MPS) 1 mold (Harley-Scheie's syndrome), MPS II mold (Hunter's syndrome), MPS IIIA mold (Sun Phi RIO A syndrome), MPS An IIIB mold (Sun Phi RIO B syndrome), MPS IIIC mold (Sun Phi RIO C syndrome), MPS An IIID mold (Sun Phi RIO D syndrome), MPS An IVA mold (MORUKIO A syndrome), MPS An IVB mold (MORUKIO B syndrome), MPS VI mold (MAROTO-rally syndrome), MPS It is a VII mold (Sly's syndrome).

[0085]

In a desirable mode, the DNA array as which Target DNA is chosen, including mutation includes the forward Tsuneno greensand-mold array which can correct mutation.

In a desirable mode, mutation has Target DNA in a cystic-fibrosis film penetration controlling factor (CFTR) gene, including mutation. Preferably, mutation changes the amino acid of the codon 508 of a CFTR protein coding region, and it is 3 base-pair in frame deletion from which mutation removes the phenylalanine of the codon 508 of CFTR protein. This deletion of the phenylalanine -508 in CFTR protein is looked at at a high rate by the test subject who has the cystic fibrosis. Therefore, in a desirable mode, a mutation CFTR gene is made into a target using the DNA array including the array which carries out the code of the phenylalanine -508 which is looked at by the wild type CFTR gene chosen, and correcting is possible.

[0086]

In a desirable mode, mutation has Target DNA in a Homo sapiens beta globin gene, including mutation. Preferably, mutation changes the amino acid of the sixth codon of a beta globin gene, and mutation is the permutation of A [ in / the sixth codon of a beta globin gene ] to T. This mutation draws the change to the valine looked at by the test subject who has sickle-cell anemia from glutamic acid in beta globin protein. Therefore, in a desirable mode, a mutation beta globin gene is made into a target using the DNA array which carries out the code of the wild type amino acid residue by the codon 6, which contains A so that it may see within the sixth codon of DNA chosen, for example, a wild type beta globin gene, and which is chosen, and correcting is possible.

[0087]

In a desirable mode, mutation has Target DNA in a factor VIII gene, including mutation. For example, mutation may be in the exons 23 and 24 of a factor VIII gene, and/or an exon 25. Preferably, mutation changes amino acid by the codon 2209 of the coding region of a factor VIII protein coding region, and it is the permutation of G in the exon 24 of the factor VIII gene to which mutation leads the change to a glutamine from an arginine from the amino acid 2209 of Factor VIII to A. Preferably, mutation changes amino acid by the codon 2229 of the coding region of a factor VIII protein coding region, and it is the permutation of G in the exon 25 of the factor VIII gene to which mutation leads the change to a cysteine from a tryptophan from the amino acid 2229 of Factor VIII to T. Such mutation has been related with serious hemophilia A from whenever [ middle ]. Therefore, in a desirable mode, a mutation factor VIII gene is made into a target using the DNA array containing DNA which carries out the code of the wild type amino acid by the codon 2209 of the coding region of a factor VIII gene, one of DNA which carry out the code of the wild type amino acid by the codon 2229 of the coding region of a factor VIII gene, or both chosen, and correcting is possible.

[0088]

In a desirable mode, mutation has Target DNA in a factor IX gene, including mutation. For example, in the test subject who has hemophilia B, the great portion of mutation is the point mutation in a factor IX gene. Therefore, in a desirable mode, the DNA array chosen makes a target one or more point mutation in the factor IX gene relevant to hemophilia B, and in order to correct, one or more nucleotides which



have at least one nucleotide of the wild type factor IX gene origin can be included.

[0089]

In a desirable mode, mutation has Target DNA in a von Willebrand factor gene, including mutation. Preferably, mutation is the single cytosine deletion under 6 cytosine expanding of 2679-2684 also in the exon 18 of the Von Will Brando gene. This mutation is looked at at a remarkable rate by the test subject who has Von Will Brando disease 3 mold. Other mutation relevant to Von Will Brando disease 3 mold, for example, point mutation, is changeable so that it may be indicated by this specification. Therefore, in a desirable mode, a mutation Von Will Brando gene is made into a target using the DNA array at least that [ whose ] of the exon 18 of the array looked at by the wild type Von Will Brando gene, for example, the Von Will Brando gene, contains six cytosines of 2679-2684 and which is chosen, and correcting is possible.

[0090]

In a desirable mode, mutation has Target DNA in a xeroderma pigmentosum group G (XP-G) gene, including mutation. Preferably, at least that of the 245 base-pair exon as which mutation is regarded by the XP-G gene is the single adenine deletion under adenine expanding of 19-21. This deletion draws a xeroderma pigmentosum. Therefore, in a desirable mode, a mutation XP-G gene is made into a target using DNA in which the wild type array of a XP-G gene, for example, that of the 245 base-pair exon of a XP-G gene, contains three adenines of 19-21 and which is chosen, and correcting is possible.

[0091]

In another desirable mode, an alteration includes the homologous recombination between the DNA array chosen and Target DNA, for example, a chromosome.

In a desirable mode, the DNA arrays chosen are a sufficient number of nucleotides, and differ from a target so that a target or the DNA array chosen may have a non-pair field, for example, a loop-formation out field, by more nucleotides than 1 unlike Target DNA. In such application, Msh2, Msh6, Msh3, Mlh1, Pms2, Mlh3, Pms(es)1, or those combination may also be offered as some complex.

[0092]

DNA as which an alteration is chosen in a desirable mode, including the nest of the array to Target DNA chosen is incorporated as it has the element on a target chosen beforehand, and the relation chosen beforehand, for example, one side is a control member, and it functions as a control member adjusting the manifestation of a protein coding sequence, when it is the array to which another side carries out the code of the protein. The contiguity array which promotes the nest chosen may be used. The DNA array chosen can be incorporated [ 5' of the target sequence chosen, for example, a gene, and a coding sequence, and ] in 3' or this array.

[0093]

In a desirable mode, the DNA array chosen including the nest of the DNA array as which an alteration is chosen is a control array, for example, an exogenous control array. In a desirable mode, a control array includes one or more :promoters and enhancers, UAS, a scaffold adhesion field, or a transcription factor binding site. In a desirable mode a control array A :metallothionein-I gene, For example, the control array of the mouse metallothionein gene origin, the control array of the SV-40 gene origin, The control array of the cytomegalovirus gene origin, the control array of the collagen gene origin, The control array of the actin gene origin, the control array of the immunoglobulin gene origin, The control array of the HMG-CoA reductase gene origin, the control array of gamma actin gene origin, the control array of the transcriptional activator YY1 gene origin, the control array of the fibronectin gene origin, or the control array of the EF-1 alpha gene origin is included.

[0094]

In a desirable mode, the DNA array chosen contains an exon. Preferably, an exogenous exon contains the code DNA of the endogenous gene made into :CAP part, a nucleotide sequence ATG, and/or a target, and an in frame.

[0095]

In a desirable mode, the DNA array chosen includes a splice-donor part.

In a desirable mode, when the DNA array chosen is incorporated in label, it includes the exogenous



control array which acts so that endogenous gene expression may be controlled. DNA chosen is set in label, and it sets in label and it can be included [ the upstream of the coding region of an endogenous gene, or ] in the upstream of the endogenous control array of an endogenous gene or a coding region. In another desirable mode, DNA chosen can be incorporated in an endogenous gene, the lower stream of a river of a coding region, the intron, or an endogenous gene. In another desirable mode, the endogenous control array of an endogenous gene is inactive, for example, deletion is carried out completely partially.

[0096]

In a desirable mode, the DNA array chosen is the upstream of an endogenous gene, and is connected with the second exon of an endogenous gene.

In a desirable mode, an endogenous gene carries out the code of :hormone, cytokine, an antigen, an antibody, an enzyme, a coagulation factor, transporter protein, an acceptor, control protein, structural protein, or the transcription factor. In a desirable mode an endogenous gene The following protein:erythropoietin, Calcitonin, a growth hormone, an insulin, in SURINO tropine, an insulinlike growth factor, Parathyroid hormone, alpha2-interferon (IFNA2), a beta interferon, A gamma interferon, a nerve growth divisor class, FSHbeta, TGF-beta, a tumor necrosis factor, Glucagon, the bone growth factor -2, the bone growth factor -7, TSH-beta, interleukin 1, interleukin 2, interleukin 3, interleukin 6, interleukin 11, interleukin 12, CSF-granulocyte (GCSF), a CSF-macrophage, CSF-granulocyte / macrophage, immunoglobulins, and catalyst nature antibodies Proteinkinase C, a glucocerebrosidase, superoxide dismutase, A tissue plasminogen activator, urokinase, Antithrombin III, DNase, the alpha-galactosidase, tyrosine hydroxylase, a blood coagulation factor V A blood coagulation factor VII, a blood coagulation factor VIII, a blood coagulation factor IX, a blood coagulation factor X A blood coagulation factor XIII, apolipoprotein E, apolipoprotein A-I, Globins, a low-density lipoprotein acceptor, IL-2 acceptor, and IL-2 antagonists An alpha-1-anti trypsin, immunoreaction qualification agents, beta-GURUKO ceramidase, alpha-IZURONIDAZE, alpha-L-IZURONIDAZE, glucosamine-N-sulfatase, Alpha-N-acetyl glucosaminidase, acetyl-coenzyme-A:alpha-glucosamine-N-acetyltransferase, The code of either N-acetyl glucosamine-6-sulfatase, the beta-galactosidase, the beta-glucuronidase, N-acetyl galactosamine-6-sulfatase and the fusibility CD 4 is carried out.

[0097]

In a desirable mode, an endogenous gene carries out the code of the follicle-stimulating hormone beta (FSHbeta), and the DNA array chosen includes the control array from which a control array, for example, the control array and array of an FSHbeta gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of an FSHbeta coding region (array number 1). In a desirable mode, a targeting array is the FSHbeta coding sequence upstream within an FSHbeta coding sequence, and directs homologous recombination. In a desirable mode, a targeting array contains the nucleotide of the array number 3 origin corresponding to -155 which 20, 30, 50,100, or 1000 follow at least from the nucleotide -696 of the nucleotide -7454 of a Homo sapiens FSHbeta array to the array number 2 origin corresponding to -1417 (numbering receives at least a translation open starting area) or a Homo sapiens FSHbeta array.

[0098]

In a desirable mode, an endogenous gene carries out the code of the interferon-alpha 2 (IFNalpha2), and the DNA array chosen includes the control array from which a control array, for example, the control array and array of IFNalpha2 gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of IFNalpha2 coding region. In a desirable mode, a targeting array directs homologous recombination in the field of the IFNalpha2 coding-sequence upstream. In a desirable mode, a targeting array contains the nucleotide of the array number 4 origin corresponding to -511 (numbering receives at least a translation open starting area) which 20, 30, 50,100, or 1000 follow at least from the nucleotide -4074 of Homo sapiens IFNalpha2 array. For example This array : From a nucleotide -4074 to at least 20, 30 and 50, or



100 nucleotides of the array number 7 origin corresponding to -3796 of Homo sapiens IFNalpha2 array; At least 20 of the array number 8 origin corresponding to -510 from the nucleotide -582 of Homo sapiens IFNalpha2 array, 30, Or 50 nucleotides; at least 20, 30, 50 and 100, or 1000 nucleotides of the array number 9 origin corresponding to -583 may also be included from the nucleotide -3795 of Homo sapiens IFNalpha2 array.

[0099]

In a desirable mode, an endogenous gene carries out the code of the granulocyte colony-stimulating factor (GCSF), and the DNA array chosen includes the control array from which a control array, for example, the control array and array of a GCSF gene, differs. one or more edges of a DNA array where a targeting array adjoins the DNA array chosen, for example, such a targeting array is preferably chosen as it -- it exists in both ends preferably. In a desirable mode, a targeting array is homologous to the field of 5' of a GCSF coding region. In a desirable mode, a targeting array is the GCSF coding sequence upstream within a GCSF coding sequence, and directs homologous recombination. In a desirable mode, a targeting array contains the nucleotide of the array number 5 origin corresponding to 101 (numbering receives at least a translation open starting area) which 20, 30, 50,100, or 1000 follow at least from the nucleotide -6,578 of a Homo sapiens GCSF array. For example, a target sequence may also contain 20, 30, 50,100, or 1000 nucleotides of the array number 6 origin corresponding to -364 (numbering receives at least a translation open starting area) from the nucleotide -6,578 of a Homo sapiens GCSF gene.

[0100]

In another desirable mode, as for a DNA array, a DNA array carries out the code of the protein, including a coding region. In a desirable mode, a coding region carries out the code of :hormone, cytokine, an antigen, an antibody, an enzyme, a coagulation factor, transporter protein, an acceptor, control protein, structural protein, or the transcription factor. In a desirable mode a coding region The following protein:erythropoietin, Calcitonin, a growth hormone, an insulin, in SURINO tropine, an insulinlike growth factor, Parathyroid hormone, a beta interferon, a gamma interferon, A nerve growth divisor class, FSHbeta, TGF-beta, a tumor necrosis factor, glucagon, The bone growth factor -2, the bone growth factor -7, TSH-beta, interleukin 1, interleukin 2, interleukin 3, interleukin 6, interleukin 11, interleukin 12, CSF-granulocyte, A CSF-macrophage, CSF-granulocyte / macrophage, and immunoglobulins Catalyst nature antibodies, proteinkinase C, a glucocerebrosidase, superoxide dismutase, A tissue plasminogen activator, urokinase, Antithrombin III, DNase, the alpha-galactosidase, tyrosine hydroxylase, a blood coagulation factor V A blood coagulation factor VII, a blood coagulation factor VIII, a blood coagulation factor IX, a blood coagulation factor X A blood coagulation factor XIII, apolipoprotein E, apolipoprotein A-I, Globins, a low-density lipoprotein acceptor, IL-2 acceptor, and IL-2 antagonists An alpha-1-anti trypsin, immunoreaction qualification agents, beta-GURUKO ceramidase, alpha-IZURONIDAZE, alpha-L-IZURONIDAZE, glucosamine-N-sulfatase, Alpha-N-acetyl glucosaminidase, acetyl-coenzyme-A:alpha-glucosamine-N-acetyltransferase, The code of either N-acetyl glucosamine-6-sulfatase, the beta-galactosidase, the beta-glucuronidase, N-acetyl galactosamine-6-sulfatase and the fusibility CD 4 is carried out.

[0101]

In a desirable mode, the DNA array chosen can be incorporated in the target of the lower stream of a river of an endogenous control array or the upstream of the coding region of an endogenous gene, and the lower stream of a river of the endogenous control array of a gene. In another desirable mode, the thing whose coding region of an endogenous gene is inactive and to include in the lower stream of a river of an endogenous control array is possible for the DNA array chosen so that deletion may be carried out, for example.

[0102]

In another side face, this invention is characterized by the cell created by either of the approaches indicated by this specification.

In another side face, this invention is characterized by the approach of changing the manifestation of the protein coding sequence of a gene in a cell with either of the approaches indicated by this specification.

[0103]



It contains maintaining a homologous recombination cell under the conditions which this approach enables the alteration of the genome array which introduces complex given [ this ] in a specification into intracellular, and is made into; target which has a DNA array including a control array in a desirable mode, and maintain a cell and enable the manifestation of the protein coding sequence of a gene under accommodation of; and a control array under the condition which produces a homologous recombination cell.

[0104]

Under accommodation of a control array, under the conditions which enable the manifestation of the protein coding sequence of a gene, a homologous recombination cell is maintained and this changes the manifestation of the protein coding sequence of a gene.

[0105]

In this specification, the vocabulary "homologous" is the same as that of a target site, for example, a chromosome DNA target site, or points out a targeting array similar enough so that a targeting array and a target site may be able to pass through homologous recombination. As long as homologous recombination can occur by useful frequency, the base pair mismatch of a small rate is permissible.

[0106]

In this specification, the vocabulary "a wild type" does not cause this disease or a functional disorder, for example [ / a disease or a functional disorder ], it does not contribute to it, but points out the array which does not influence or adjust it.

[0107]

In this specification, "complex" points out the stable meeting to which coupling of the component is carried out by a share or the noncovalent bond.

Probably, other descriptions and advantages of this invention will be clear from the following detailed explanation to a claim.

Detailed description \*\* which increases homologous recombination Since the homologous recombination between the DNA array chosen and Target DNA DNA, for example, a chromosome, is promoted, it is possible to offer \*\* which increases homologous recombination with the DNA array chosen. \*\* which increases homologous recombination The effectiveness of the chain invasion during the function;3 recombinant-DNA array to which the homologous pair formation between the parts chosen as the DNA array and nest to which the homologous recognition between the parts chosen as the DNA array and nest of the one or more followings by which function, : 1 selection are made is made to increase, and by which function;2 selection is made is made to increase, and chain exchange It has the function to which the processing effectiveness of the middle structure to the mature product of the function;4 recombination to which it is made to increase is made to increase.

[0108]

In mixture including a double-stranded-DNA array, can introduce into a cell, and can introduce just before administration of a DNA array, or into immediately after, it is made to adhere on a DNA array, for example, \*\* which increases homologous recombination can be coated. It is \*\* 52 52 which increases homologous recombination, for example, Rad, for example, hRad, and its fragment, and it is possible to coat all DNA arrays, it is possible to coat one or more edges of a DNA array, for example, it is possible to coat one or more protrusion single strand edges of a DNA array. Preferably, \*\* which increases homologous recombination coats a part, even if there are protrusion single strand 3 'edge or 5' few edges of a DNA array.

[0109]

the example of \*\* which increases homologous recombination -- : -- Rad52 or functioning fragment; -- Rad51 or functioning fragment; -- two or more combination of the fragment of Rad54, functioning fragment; these protein, or these protein is included. Intracellular can be discovered and \*\* which increases homologous recombination can introduce into a cell the nucleic-acid array which carries out the code of either of the above-mentioned \*\* again, for example.

[0110]

It is possible to make a decision with Rad51 functional fragment with a known technique. For example,



the functionality of Rad51 fragment is [ for example, / Baumann / Cell ] (1996). in known by the technical field concerned as indicated in 87:757-766 In vitro assay, it can determine based on the capacity to mediate homologous pair formation and chain exchange. Briefly, preincubation of the hRad51 is first carried out to annular ssDNA, and 32P indicator straight-line duplex deoxyribonucleic acid is added after that. Electrophoresis can determine the amount of formation of a connection molecule, and chain exchange. Furthermore, the functionality of Rad51 fragment is [ Benson / EMBO ] (1994). J. As indicated in 13:5764-5771, under existence of ATP, it can combine with the duplex chain by which nick formation was carried out, and can determine based on the capacity which forms the whorl nucleoprotein filament which can be visualized with an electron microscope. In the cell which lacks the functioning Rad51 protein again, even if the functionality of Rad51 is based on the capacity which mitigates DNA repair and the defect in homologous recombination, it can be determined. Therefore, when it compares with the bottom of the nonexistence, if it has effect of electropositive in above-mentioned assay, it can be determined whether Rad51 fragment functions. Furthermore, the degree of the effect of electropositive which it has with Rad51 fragment can be compared with the degree of the effect of electropositive which it has with an overall length Rad51.

[0111]

The functionality of Rad54 fragment is the assay known by the technical field concerned, for example, SwagemakersJ(1998). Biol. Chem. In the assay indicated by 273:28292-28297, it can determine based on the capacity which hydrolyzes ATP under existence of dsDNA. Furthermore, the functionality of Rad54 fragment can be determined in the cell lacking in the functioning Rad54 protein based on the capacity which mitigates DNA repair and the defect in homologous recombination.

[0112]

Rad52 and its functioning fragment Rad52 offered with the DNA array of the part in Target DNA chosen, for example, the part in Chromosome DNA chosen, can offer the higher alteration of the part of a rate which will be produced under the nonexistence, for example, homologous recombination. Although to be restricted by the theory is not desirable, Rad52 from the function;2 nuclease digestion which protects all DNA arrays from the following function:1 nuclease digestions The protrusion single strand edge of a DNA array, For example, the function to which the homologous recognition between the parts chosen for the function;3DNA array and nest which protect 3' tails is made to increase; it is thought possible to offer one or more [ of the function to which the homologous pair formation between the parts chosen as a list for 4DNA array and a nest is made to increase ].

[0113]

Rad(s)52 are some approaches including isolation of Rad52, or the manifestation of the coding sequence by the genetic manipulation method, and can be obtained. For example, Van Dyke(s) (1999) Nature 398:728 indicates production and purification of hRad52 from Sf9 cell. The nucleotide sequence of Rad52 of various kinds is known. For example, Shen(s) (1995) Genomics 25(1):199-206(rat and Homo sapiens Rad52);MurisMutat(s)(1994). Res. 315(3):295-305 (a rat and Homo sapiens Rad52);P arkJ (1995). Biol. Chem. Please refer to 270 (26):15467-15470 (Homo sapiens Rad52).

[0114]

or the fragment of Rad52 depends Rad52 or its part on the manifestation of the array which carries out a code -- or gene activation -- depending (desirable approach) -- it is based on proteolysis-digestion, or it is some approaches by chemical composition, and can produce. The interior or the end fragment of Rad52 can be generated by removing the end (an end fragment sake) of the nucleic acid which carries out the code of Rad52, or one or more nucleotides of the both-ends (internal fragment sake) origin. The manifestation of Mutagenesis DNA produces a Rad52 polypeptide fragment. Therefore, digestion with "end \*\*\*\* picking (end-nibbling)" endonuclease or various restriction enzymes can generate DNA which carries out the code of the array of Rad52 fragment. DNA which carries out the code of the fragment of Rad52 protein can be generated again also with the combination of the approach discussed to random shear and limit digestion or a top.

[0115]

Rad52 fragment can be chemically compounded again using the technique known by the technical fields



concerned, such as idiomatic MERIFIRUDO solid phase f-Moc or a t-Boc chemical reaction. For example, it divides into the fragment of the desirable die length which does not include duplication of a fragment at arbitration, or Rad52 peptide can be divided into the duplication fragment of desirable die length.

[0116]

It is possible to make the decision of whether Rad52 fragment functions with a known technique. For example, since it determines whether Rad52 fragment is able to protect to a nuclease digestion, it is possible to carry out the incubation of an end indicator straight-line-sized double-stranded-DNA array, for example, the <sup>32</sup>P indicator straight-line-sized double-stranded-DNA array, to Rad52 fragment before installation of nuclease, for example, exonuclease, or endonuclease. Then, it is possible to determine the emitted indicator, for example, the amount of <sup>32</sup>P. The amount of the emitted indicator serves as an index of the capacity which Rad52 fragment protects to a nuclease digestion. Furthermore, the functionality of Rad52 fragment can be determined based on the capacity which stimulates connection molecule formation. The functionality of Rad52 fragment is [ Benson / Nature ] (1998). By stimulus of hRad51 drive connection molecule formation which is indicated by 391:401-404, it is in. Analysis is possible at vitro. Briefly, preincubation of the hRad51 is first carried out to annular ssDNA, and <sup>32</sup>P indicator straight-line duplex deoxyribonucleic acid is added after that. Electrophoresis can determine formation of a connection molecule. When addition of Rad52 is compared with the connection molecule formation under the nonexistence of Rad52, it stimulates formation of a connection molecule. Therefore, if connection molecule formation is stimulated when it compares with the connection molecule formation under nonexistence, it can be determined whether Rad52 fragment functions. Furthermore, the degree of the stimulus by Rad52 fragment can be compared with the degree of overall-length Rad52 stimulus. Furthermore, the functionality of Rad52 fragment is Park(1995) J. Biol. Chem. As indicated in 270:15467-15470, when carrying out a superfluous manifestation in a culture ape cell, the resistance over ionizing radiation is made to increase and it can determine based on the capacity to which the rate of homologous recombination is made to increase.

[0117]

\*\* which checks non-homologous edge connection It is the higher rate which will be produced under nonexistence using \*\* which checks non-homologous edge connection, and it is possible to provide the part in Target DNA chosen with a DNA array. Non-homologous edge connection may draw the inaccurate fusion between double strand edges, for example, a re-connection edge may have insertion or deletion. \*\* which checks non-homologous edge connection may be what kind of \*\* which checks the manifestation and/or activity of a molecule which participate in a non-homologous edge connection path. For example, the complex of Mre11, Rad50, and Nbs1 participates in non-homologous edge connection. therefore -- for example, the thing for which formation of this complex is checked -- for example, it is possible by combining with either of these protein, or checking one manifestation of these protein to check non-homologous edge connection. Furthermore, Ku protein, Ku70 or Ku80, a ligase 4 (Lig4), and Xrcc4 are contained in other protein which participates in non-homologous edge connection. [ for example, ]

[0118]

Ku deactivator It is the part where it is chosen in Target DNA, for example, the part where it is chosen in Chromosome DNA, and if Ku deactivator is offered with a DNA array, it is possible to offer the higher alteration of the part of a rate which will be produced under the nonexistence, for example, homologous recombination. Ku is a heterodimer of about 70 kDa(s) and 80kDa(s) combined with a DNA break point, and plays a role in the double strand cutting restoration by non-homologous edge connection. "Ku80" may also be called "Ku86" again.

[0119]

Ku deactivator can check Ku manifestation or Ku activity. Preferably, it interacts with the nucleotide sequence which carries out the code of Ku or the Ku, for example, combines with it, and Ku deactivator checks Ku manifestation or Ku activity. Preferably, Ku dependence non-homologous edge connection is checked. Ku inhibitor can check Ku70, Ku80, or both.



[0120]

An antisense Ku nucleic-acid molecule is contained in anti-Ku antibody and Ku tie molecules, for example, the random generation peptide combined with Ku, Ku joint oligomer and a polymer, and a list at \*\* usable although Ku is inactivated. Preferably, \*\* which inactivates Ku is \*\* in which partial administration of anti-Ku antibody, Ku tie molecules, etc. is possible, for example, the random generation peptide combined with Ku, Ku joint oligomer, or a polymer.

[0121]

Preferably, Ku deactivator interacts with Ku, for example, is combined. \*\* which interacts with Ku protein is an alteration part, and can inactivate Ku locally.

[0122]

For example, DNA made into a DNA array and a target is approached very much, and Ku deactivator is introduced into a cell, and this checks Ku locally by the homologous recombination part. In mixture including a double-stranded-DNA array, it can introduce into a cell, and can introduce just before administration of a DNA array, or into immediately after, or Ku deactivator can carry out covalent bond to the protein 52 relevant to a DNA array or a DNA array, for example, Rad, and its fragment. A cell can carry out preincubation to Ku deactivators, such as anti-Ku antibody or an antisense Ku nucleic-acid molecule, again.

[0123]

Anti-Ku antibody It is made to combine with Ku using anti-Ku antibody or its fragment, and it is possible to decrease Ku activity by that cause. Although anti-Ku antibody is an alteration part and it interacts with Ku locally, it is possible to prescribe a medicine for the patient so that Ku manifestation generally may not be checked in a cell. Anti-Ku70 and anti-Ku80 antibody is contained in anti-Ku antibody.

[0124]

It is possible to use Ku protein, its part, or a fragment as immunogen, and to generate the antibody combined with Ku using the standard technique for a polyclonal and monoclonal antibody preparation. Overall-length Ku protein is usable, or the antigenic peptide fragment of Ku is usable as immunogen.

[0125]

Typically, an antibody is prepared using Ku or Ku peptide by carrying out immunity of the suitable test subject (for example, a rabbit, a goat, a mouse, or other mammals) by immunogen. Ku protein obtained by the manifestation of the array which carries out the code of the Ku, or gene activation, or Ku peptide compounded chemically can be included by suitable immunogenicity preparation. For example, U.S. Pat. No. 5,460,959; and connection U.S. application USSN by which the whole is clearly used for this specification 08/334,797;USSN 08/231,439;USSN 08/334,455;, and USSN Please refer to 08/928,881. The nucleotide and amino acid sequence of Ku are known, and he is Takiguchi et al. (1996). Genomics It is indicated by 35(1):129-135. Adjuvants, such as completeness or Freund's incomplete adjuvant, or the same immunity stimulative agent of Freund can be further included by preparation. The immunity of the suitable test subject using immunogenicity Ku preparation guides polyclonal anti-Ku antibody reaction.

[0126]

Anti-Ku antibody or its fragment is usable as a Ku deactivator. F (v), Fab, Fab', and F(ab')<sub>2</sub> fragment generable by processing an antibody with enzymes, such as a pepsin, is contained in the example of anti-Ku antibody fragment. The vocabulary "a monoclonal antibody" or a "monoclonal antibody constituent" points out an antibody molecule ensemble only including one kind of the antigen binding site which can carry out an immunoreaction to the specific epitope of Ku in this specification. Therefore, typically, a monoclonal antibody constituent shows the single binding affinity to specific Ku protein which carries out an immunoreaction.

[0127]

Furthermore, anti-Ku antibody produced by hereditary operation information, such as a chimera, a hominization monoclonal antibody, etc. which contain both the Homo sapiens and the nonhuman part which can be created using standard recombinant DNA technology, is usable. Such a chimera and a



hominization monoclonal antibody The standard DNA technique et al. known by the technical field concerned, for example, Robinson, the [ international application ] -- No. PCT/US86/02269; -- Akira et al. and Europe patent application 184,187; Taniguchi M. -- Morrison et al. and Europe patent application 171,496; patent application [ Europe ] No. 173,494; -- Neuberger et al. -- the [ PCT international official report ] -- WO No. 86/01533; -- Cabilly et al. -- U.S. Pat. No. 4,816,567; Cabilly et al., Europe patent application 125,023; -- Better and others -- Science 240:1041-1043, 1988; -- Liu et al. PNAS 84 :3439-3443 1987; -- Liu and others -- J. Immunol. 139:3521-3526 1987; -- Sun et al. PNAS 84:214-218 1987; -- Nishimura et al. Canc. Res. 47:999-1005 -- 1987; -- Wood et al. Nature 314:446-449 1985; and Shaw et al. J. Natl. Cancer Inst. 80:1553-1559 1988; Morrison S.L. Science 229:1202-1207 1985; -- Oi et al. BioTechniques 4:214 It Winter(s). 1986; -- U.S. Pat. No. 5,225,539; Jones and others, Nature 321:552-525, 1986; -- Verhoeyan et al. Science 239:1534, 1988; and Beidler et al. J. Immunol. 141:4053-4060 By the genetic manipulation using the approach indicated by 1988 It can produce.

[0128]

Furthermore, the Homo sapiens monoclonal antibody turned to Ku can be created using a standard technique. For example, a Homo sapiens monoclonal antibody is generable in the immune disorder mouse which transplanted the transgenic mouse or the antibody production human cell. The approach of generating such a mouse is Wood et al. and the PCT official report WO. 91/00906, Kucherlapati et al., PCT official report WO 91/10741; -- Lonberg et al. and PCT official report WO 92/03918; -- Kay et al. - PCT official report WO 92/03917; -- Kay et al. and PCT official report WO 93/12227; -- Kay et al. -- PCT official report 94/25585; -- Rajewsky et al. and PCT official report WO 94/04667; -- Ditullio et al. -- PCT official report WO 95/17085; Lonberg, N. et al. (1994) Nature 368:856-859; Green L.L. et al. (1994) Nature Genet. 7:13-21; Morrison S.L. et al. (1994) Proc. Natl. Acad. Sci. USA 81:6851-6855; -- Bruggeman et al. (1993) Year Immunol 7:33-40; -- Choi et al. (1993) Nature Genet. (1993) [ 4:117-123; Tuailon ] PNAS 90:3720-3724; -- Bruggeman et al. (1991) Eur J Immunol 21:1323-1326; -- Duchosal et al. -- PCT official report WO 93 / 05796; U.S. Pat. No. 5,411,749; -- McCune et al. (1988) Science 241:1632-1639, Kamel-Reid et al. (1988) Science 242:1706; Spanopoulou (1994) Genes & Development 8:1030-1042; -- Shinkai et al. (1992) Cell It is indicated by 68:855-868. Immunity of the immune disorder mouse which transplanted the Homo sapiens antibody-transgenic mouse, the Homo sapiens antibody forming cell, or the organization is carried out with Ku or an antigenic Ku peptide, and it is possible to generate a hybridoma after that using the spleen cell of these immunity mouse origins. The hybridoma producing method is well-known.

[0129]

The Homo sapiens monoclonal antibody to Ku can be prepared again using the immunoglobulin light chain and heavy chain cDNA which are prepared from mRNA of a test subject's lymphocyte origin also by building combinatorial immunoglobulin libraries, such as a Fab phage display library or a scFv phage display library. for example, McCafferty et al. and the PCT official report WO 92/01047; -- Marks et al. (1991) J. Mol. Biol. 222:581-597; and Griffiths et al. (1993) EMBO J Please refer to 12:725-734. Furthermore, the combinatorial library of an antibody variable region is generable by carrying out mutation of the known Homo sapiens antibody. For example, it is possible to screen whether this is combined with Ku after that by carrying out mutation by generating the library of a mutation variable region using the mutagenesis oligonucleotide changed by random in the variable region of a Homo sapiens antibody where combining with Ku is known. The approach of forming opposite formation and a screening procedure combining the approach of guiding random mutagenesis, the randomized heavy chain, and a light chain, in the CDR field of an immunoglobulin heavy chain and/or a light chain is Barbas et al. and the PCT official report WO. 96/07754; Barbas Proc(s)(1992). Nat'l Acad. Sci. USA It is possible to find out to 89:4457-4461.

[0130]

It is preferably made discovered by the display package ensemble of the filamentous phage origin, and an immunoglobulin library can form an antibody display library. The approach of being easy to use it especially for generating an antibody display library, and the example of a reagent Ladner et al., U.S. Pat. No. 5,223,409; Kang et al. [ for example, ], PCT official report WO 92/18619; -- Dower et al. and



PCT official report WO 91/17271; -- Winter et al. -- PCT official report WO 92/20791; -- Markland et al. and PCT official report WO 92/15679; -- Breitling et al. -- PCT official report WO 93/01288; -- McCafferty et al. and PCT official report WO 92/01047; -- Garrard et al. -- PCT official report WO 92/09690; -- Ladner et al. -- PCT official report WO 90/02809; -- Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; -- Hay et al. (1992) *Hum Antibod Hybridomas* (1989) [ 3:81-85;Huse ] *Science* 246:1275-1281; -- Griffiths et al. (1993) above-mentioned; -- Hawkins et al. (1992) *J Mol Biol* (1991) [ 226:889-896;Clackson ] *Nature* 352:624-628;Gram(s) (1992) *PNAS* 89:3576-3580; -- Garrard et al. (1991) *Bio/Technology* (1991) [ 9:1373-1377;Hoogenboom ] *Nuc Acid Res* 19:4133-4137;, and Barbas et al. (1991) *PNAS* It is possible to find out to 88:7978-7982. If displayed on a display package (for example, filamentous phage) front face, an antibody library will be screened, and the package which discovers the antibody combined with Ku will be identified and isolated. In a desirable mode, the display package which discovers the antibody combined with Immobilization Ku is chosen with carrying out panning of the primary screening of a library using Immobilization Ku.

[0131]

The monoclonal antibody to Ku is commercially available also from Neomarkers (California FUREMONT) again.

Ku tie molecules The molecule combined with Ku(s), such as Ku joint peptide, for example, a random generation peptide, and Ku joint oligomer, or a polymer, is usable as a Ku deactivator. It combines with Ku protein and such a molecule can check at least one activity of Ku(s), such as non-homologous edge connection.

[0132]

The example of Ku joint oligomer is WO by which the contents are used for this specification. It is shown in 99/33971. Such oligomer can consist of a nucleotide, a nucleotide analog (analog), or combination. Preferably, oligomer consists of ribonucleotides. It is possible to identify the protein which is combined with Ku or interacts with Ku using these Ku oligomer. The approach of identifying Ku joint peptide using these oligomer is WO. It is indicated to 99/33971.

[0133]

Furthermore, it is possible to screen a random generation peptide about the capacity combined with Ku. For example, the various techniques which screen the generated mutation gene product are known by the technical field concerned. The technique which screens a big gene library often carries out cloning of the gene library into [ which can be reproduced ] an expression vector, and includes making a gene discover under the conditions which promote comparatively easy isolation of the vector which is the produced vector library, carries out the transformation of the suitable cell, and carries out the code of the gene in which detection of association to desirable activity, for example, Ku, detects the product. Each of the technique indicated below is accepted in the high processing analysis for screening the array of generated a large number for example, with a random mutagenesis technique.

[0134]

Display library In another approach which screens Ku joint peptide, a candidate peptide is displayed on a cell or a virion front face, and the capacity combined with Ku protein through the product with which a specific cell or virion was displayed is detected by "panning assay." For example, it is possible to carry out cloning of the gene library into the gene of the surface membrane protein of a bacterial cell, and to detect the produced fusion protein by panning (1371;, and Goward et al. [ Ladner et al. / WO 88/06630;Fuchs s(1991) *Bio/Technology* 9:1370- ] (1992) *TIBS* 18:136 -140). Score attachment is possible about the peptide homologue which functions potentially by the same method using detectable indicator ligand. The homologue holding ligand avidity is detectable using fluorescent-labeling ligand. If fluorescent-labeling ligand is used, when it will become possible under a fluorescence microscope to inspect a cell visually and to dissociate or the gestalt of a cell will allow, a fluorescence activation cell aliquot enables it to dissociate.

[0135]

As for a gene library, it is possible to make it discovered as fusion protein on a virion front face. For example, in a filamentous phage system, it is possible to make a heterogeneous (foreign) peptide array



discover on the front face of infective phage, and for this to give two significant advantages. First, these phage is the concentration per ml and far exceeding 10<sup>13</sup> phage, and since it is applicable to an affinity matrix, it can screen much phage at once. To it, each infective phage can amplify the phage with another infection period, if specific phage is collected from an affinity matrix with low yield by the second, since a gene product is displayed on the front face. The almost same *Escherichia coli* (*E. coli*) filamentous phage and the group of M13, fd, and f1 are most frequently used in a phage display library. It is possible to generate fusion protein, without disturbing the final packaging of virion using Phage gIII or one of gVIII coat protein. Make a heterogeneous epitope discover at the NH<sub>2</sub> end of pIII, and phage with such an epitope it is possible to collect from a lot of superfluous phage lacking in this epitope (Ladner et al. --) PCT official report WO 90/02909; -- Garrard et al. -- PCT official report WO 92/09690; -- Marks et al. (1992) *J. Biol. Chem.* 267:16007-16010; -- Griffiths et al. (1993) *EMBO J* (1991) [ 12:725-734;Clackson ] *Nature* 352:624-628;, and Barbas et al. (1992) *PNAS* 89:4457-4461. [0136]

General approach uses the maltose acceptor (adventitia protein, LamB) of *Escherichia coli* as a peptide fusion partner (5 Charbit et al. (1986) *EMBO* 3029 -3037). An oligonucleotide is inserted in the plasmid which carries out the code of the LamB gene, and the peptide united with one of the proteinic extracellular loop formations is produced. These peptides are available to association to ligand, for example, an antibody, and in case they medicate an animal with a cell, they can pull out an immunoreaction. A huge bacteria surface structure is used as a vehicle for a peptide display with other cell surface protein (91 Schorr et al. (1991) *Vaccines* pp. 387 -392), for example, OmpA, PhoE (88 Agterberg et al. (1990) *Gene* 37 -45), and PAL (9 Fuchs et al. (1991) *Bio/Tech* 1369 -1372). A polymerization is carried out and they are the cilia for the exchange between bacteria of gene information. - A peptide may be united with the pilin which is protein which forms a conduit (55 Thiry et al. (1989) *Appl. Environ. Microbiol.* 984 -993). Cilia provide presentation of the peptide to an extracellular environment with a useful base material for the role at the time of interacting with other cells. Another huge surface structures used for a peptide display are a bacteria transportation organ and a flagellum. Fusion of the peptide to a protein subunit and flagellin offers the high density array of many peptide copies on a host cell (6 Kuwajima et al. (1988) *Bio/Tech.* 1080 -1083). The surface protein of other bacteria kinds is also used as a peptide fusion partner. *Staphylococcus* group (*Staphylococcus*) protein A and the adventitia protease IgA of *Neisseria* (*Neisseria*) are contained in an example (9 -4245, and Klauser et al. [ 174 Hansson et al. (1992) *J. Bacteriol.* 4239 ] (1990) *EMBO J.* 1991 -1999). [0137]

In an above-mentioned filamentous phage system and an above-mentioned LamB system, the physical connection between a peptide and its code DNA takes place by including DNA in the particle (a cell or phage) which possesses a peptide on the front face. Peptide prehension catches DNA of a particle and the interior. Another plan forms connection between a peptide and DNA using DNA-binding protein and LacI (Cull et al. (1992) *PNAS USA* 89:1865 -1869). The plasmid containing the LacI gene which has an oligonucleotide cloning part in 3' edge is used for this system. LacI peptide fusion protein is produced under the accommodation induction by arabinose. This fusion holds the natural capacity combined with the short DNA array in which LacI is known as a LacO operator (LacO). By having LacO of 2 copies on a manifestation plasmid, LacI-peptide fusant is closely combined with the plasmid which carries out the code of this. In order that each intracellular plasmid may discover only a peptide array with each single cell, including only a single oligonucleotide array, a peptide meets to stability specifically with the DNA array which directs the composition. The cell of a library is dissolved quietly and the complex which exposes peptide-DNA complex to the matrix of a fixed acceptor, and contains biologically active peptide is collected. Then, meeting plasmid DNA is again introduced into a cell for magnification, and sequencing of the DNA is carried out, and the identity of peptide ligand is determined. The big random library of a dodeca peptide was created as proof of the practical usefulness of an approach, and it chose on the monoclonal antibody created to opioid peptide and Dynorphine B. The cohorts of a peptide altogether related according to the consensus sequence corresponding to 6 residue parts of Dynorphine B were collected (Cull et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89 -1869).



[0138]

By the way, this plan is called a peptide-ON-plasmid and differs from the phage displaying method at two important points. First, a peptide adheres to the C terminal of fusion protein, and produces the display of a library member as a peptide which has an uncombined (free) carboxy end. A guest peptide is placed into the amino terminal domain which filamentous phage coat protein, pIII, and pVIII are both moored to phage through the C terminal, and is elongated on; and the outside. the peptide displayed by phage in some designs -- fusion protein -- it is exactly shown by the amino terminus (87 Cwirla et al. (1990) Proc. Natl. Acad. Sci. U.S.A. 6378 -6382). The second difference is biological bias of a lot which affects the peptide ensemble who actually exists in a library. A LacI fusion molecule is restricted to the cytoplasm of a host cell. The amino terminal which contains a peptide while it was quickly secreted by the periplasm partition through intima and the C terminal hydrophobic domain had moored to the film, although phage coat fusant was exposed to a short time and cytoplasm during the translation is projected to the periplasm, while waiting for the assembly to a phage particle. The peptides in LacI and a phage library may differ intentionally as a result of exposure to different proteolysis activity. Phage coat protein needs the transportation and signal peptidase processing to which intima is crossed as the opening scene [ phage ] of incorporation. A specific peptide demonstrates effect harmful to these processes, and too little presentation is carried out all over a library (Gallop et al. (1994) J. Med. Chem. 37 (9) : 1233 -1251). Such specific bias is not the factors of a LacI display system.

[0139]

The number of small peptides available to a recombination random library is immense. The library of the independent clone of 107-109 is prepared daily. Although the library of magnitude comparable as 1011 recombinant has been generated, this magnitude is approaching the practical limitation of a clone library. This limitation of library size happens at the process which carries out the transformation of the DNA containing a randomization part to a host bacterial cell. in based on the display of the new peptide in polysome complex in order to avoid this limitation The vitro system has been developed recently. This display library method has possibility of producing a library larger 3 to 6 figures than phage / phagemid, or a plasmid library available now. Furthermore, construction of a library, manifestation of a peptide, and screening are completely performed by cell non-\*\*\*\*\*.

[0140]

They are imprint/translation system and in about the library which built the molecule DNA library which carries out the code of the 1012deca peptide in one application of this approach (Gallop et al. (1994) J. Med. Chem. 37 (9) : 1233 -1251), and was discovered with Escherichia coli S30. vitro coupling was carried out. The complex containing the new peptide which chose conditions so that ribosome might be kept back on mRNA, was made to accumulate RNA of a remarkable rate into a polysome, and has still been connected with Code RNA was produced. A polysome is the almost same method as screening a more idiomatic recombination peptide display library, and although affinity purification is carried out on a fixed acceptor, it is fully strong. RNA of the junctional complex origin is collected, and it changes into cDNA, and it amplifies by PCR, and the template of the next period of composition and screening is produced. Coupling of the polysome displaying method may be carried out to a phage display system. Cloning of the cDNA of the concentration pool origin of a polysome was carried out to the phagemid vector after screening of several round term. This vector serves as a DNA sequencing vector for peptide identification again as a peptide expression vector which displays the peptide united with coat protein. By making a polysome origin peptide discover on phage, it is possible to continue the compatibility sorting by selection of this format, or it is possible to carry out assay of the peptide on each clone about the joint singularity in completion phage ELISA (204 Barret et al. (1992) Anal. Biochem 357 -364), concerning the avidity in Phage ELISA. In order to identify the array of biologically active peptide, sequencing of the DNA produced by the phagemid host is carried out.

[0141]

Antisense Ku nucleic-acid array The nucleic-acid molecule which is antisense to the nucleotide which carries out the code of the Ku is usable as a deactivator which checks Ku manifestation. It is complementary to "sense" nucleic acid which carries out the code of the Ku to an "antisense" nucleic



acid, for example, it is complementary to the code chain of a double strand cDNA molecule, or a complementary nucleotide sequence is included in a mRNA array. Therefore, antisense nucleic acid can form a sense nucleic acid and hydrogen bond. Antisense nucleic acid may be complementary to all the Ku code chains, or may be complementary only into the part. For example, the antisense nucleus acid-content child who is antisense is usable to the "coding region" of the code chain of the nucleotide sequence which carries out the code of the Ku.

[0142]

The code chain array which carries out the code of the Ku is Takiguchi et al. (1996). Genomics It is possible to design antisense nucleic acid according to the regulation of Watson and click base pair formation in consideration of being indicated by 35(1):129-135 and the Genbank deposition number L35932. An antisense nucleus acid-content child is Ku. Although it may be complementary to all the coding regions of mRNA, it is Ku more preferably. It is the oligonucleotide which is antisense only to the code of mRNA, or the part of a non-coding region. For example, an antisense oligonucleotide is Ku. It may be complementary to the field which surround at least the translation open starting area of mRNA. Antisense oligonucleotides may be die length 5 [ about ], 10, 15, 20, 25, 30, 35, 40 and 45, or 50 nucleotides. Antisense nucleic acid can be built using the chemosynthesis and enzyme connection using the approach learned by the technical field concerned. For example, antisense nucleic acid (for example, antisense oligonucleotide) can be chemically compounded using the various qualification nucleotides designed so that the physical stability of the duplex chain which is made to increase the biological stability of a molecule or is formed between antisense one and a sense nucleic acid might be made to increase using a natural existence nucleotide, for example, a phosphorothioate derivative and its acridine permutation nucleotide are usable. For the example of a qualification nucleotide usable although antisense nucleic acid is generated 5-fluorouracil, the 5-bromouracil, 5-chloro uracil, 5-iodine uracil, hypoxanthine, a xanthin, 4-acetyl cytosine, 5-(carboxy hydroxymethyl) uracil, 5-carboxymethyl aminomethyl-2-thiouridine, 5-carboxymethyl aminomethyl uracil, a dihydrouracil, beta D-galactosyl KEOSHIN, Inosine, N6-isopentenyladenine, 1-methyl guanine, 1-methylinosine, 2 and 2-dimethyl guanine, 2-methyladenine, 2-methyl guanine, 3-methylcytosine, 5-methylcytosine, an N6-adenine, 7-methyl guanine, 5-methylamino methyl uracil, the 5-methoxy aminomethyl-2-thiouracil, Beta D-MANNOSHIRUKEOSHIN, a 5'-methoxy carboxymethyl uracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, A uracil-5-oxy-acetic acid (v), WAIBUTOKISOSHIN, a PUSOIDO uracil, KEOSHIN, 2-thiocytosine, the 5-methyl-2-thiouracil, the 2-thiouracil, 4-thiouracil, 5-methyluracil, and uracil-5-oxy-acetic-acid methyl ester, The uracil-5-oxy-(acetic-acid v) 5-methyl-2-thiouracil, 3-(3-friend no 3-N-2-carboxypropyl) uracil, w (acp3), and 2,6-diaminopurine are contained. Or a nucleic acid can produce antisense nucleic acid biologically using the expression vector by which subcloning is carried out in the antisense direction (that is, RNA imprinted from the insertion nucleic acid is the thing of the antisense direction to the target target-nucleus acid).

[0143]

Exogenous DNA array As for the DNA array which it is going to offer for example, introduce into a cell, in a cell, it is possible to make a target sequence change. For example, it is possible to introduce a different DNA array from Target DNA chosen by 3 or less than 2 [ 10, 8, 5, 4, ], one nucleotide, for example, a permutation, deletion, or insertion. The DNA array which the DNA array chosen may differ from more nucleotides than 1 and a target sequence again, for example, is chosen differs from some nucleotides and a target sequence so that it may have a non-pair field, for example, a loop-formation out field. These alterations can embellish a target sequence manifestation. The array manifestation embellished activates the array which is usually a silence (it is not discovered), for example, a code DNA array, for example, the coding sequence usually looked at by the cell, in :cell, and sets it into; cell. Make the manifestation of the array discovered lower than normal level, for example, a code DNA array, for example, the coding sequence usually looked at by the cell, increase, and it sets into; cell. Usually, so that the array discovered with a deficit mold, for example, a code DNA array, for example, the coding sequence usually looked at by the cell, may be made to discover and it may differ from the usual pattern of; cell It includes decreasing the manifestation of the coding sequence which control or



the induction pattern of an array, for example, a code DNA array, for example, the coding sequence usually looked at by the cell, is changed, and is usually looked at by; array, for example, a code DNA array, for example, a cell, from the usual manifestation level in a cell.

[0144]

It is possible to introduce a different DNA array from Target DNA chosen by 3 or less than 2 [ 10, 8, 5, 4, ], one nucleotide, for example, a permutation, deletion, or insertion. for example, the array made into a target -- 3 or less than 2 [ a wild type array, 10, 8, 5 and 4, and ] -- or it may differ 1 nucleotide.

Preferably, the array made into a target differs from a wild type array by the mutation produced from point mutation, for example, insertion, deletion, or a permutation. Preferably, the mutation in a target sequence, for example, a gene, adjusts this disease or a functional disorder in relation to a disease or a functional disorder. Although mutation, for example, point mutation, is not necessarily limited to the example of the gene related with the disease or the functional disorder, a cystic-fibrosis film penetration controlling factor (CFTR) gene, a beta globin gene, a factor VIII gene, a factor IX gene, a von Willebrand factor gene, and a xeroderma pigmentosum group G (XP-G) gene are contained. In order to make a target sequence change, the DNA array chosen may also include the forward Tsuneno greensand-mold array which can correct mutation. According to the approach indicated by this specification, there are some changeable genetic defects and genes.

[0145]

In another side face, the DNA array which the DNA array chosen may differ from more nucleotides than 1 and a target sequence again, for example, is chosen differs from some nucleotides and a target sequence so that it may have a non-pair field, for example, a loop-formation out field. For example, when homologous recombination can be carried out with the label element chosen beforehand by the DNA array chosen, for example, one side is a control member and it is the array to which another side carries out the code of the protein, a control member controls the manifestation of a protein coding sequence. The DNA array chosen may be a control array, for example, an exogenous control array. A control array includes a promotor, an enhancer, UAS, a scaffold adhesion field, and a transcription factor binding site. Furthermore, the code DNA of an exon, the intron, a CAP part, nucleotide sequence ATG, a marker, for example, a selective marker, a splice-donor part and/or a target sequence, and an in frame can be included by the DNA array chosen again. The DNA array chosen can also include the DNA array which carries out the code of a coding region, for example, the protein, again.

[0146]

A coding sequence may be endogenous, for example, the DNA array which the DNA array chosen is a control array, or is chosen may also contain a coding region, namely, a coding region is exogenism. The coding region may be carrying out the code of the various protein. the example of such protein -- : -- erythropoietin, calcitonin, and a growth hormone -- An insulin, in SURINO tropine, an insulinlike growth factor, parathyroid hormone, alpha2-interferon (IFNA2), a beta interferon, a gamma interferon, A nerve growth divisor class, FSHbeta, TGF-beta, a tumor necrosis factor, glucagon, The bone growth factor -2, the bone growth factor -7, TSH-beta, interleukin 1, interleukin 2, interleukin 3, interleukin 6, interleukin 11, interleukin 12, CSF-granulocyte (GCSF), A CSF-macrophage, CSF-granulocyte / macrophage, and immunoglobulins Catalyst nature antibodies, proteinkinase C, a glucocerebrosidase, superoxide dismutase, A tissue plasminogen activator, urokinase, Antithrombin III, DNase, the alpha-galactosidase, tyrosine hydroxylase, a blood coagulation factor V A blood coagulation factor VII, a blood coagulation factor VIII, a blood coagulation factor IX, a blood coagulation factor X A blood coagulation factor XIII, apolipoprotein E, apolipoprotein A-I, Globins, a low-density lipoprotein acceptor, IL-2 acceptor, and IL-2 antagonists An alpha-1-anti trypsin, immunoreaction qualification agents, beta-GURUKO ceramidase, alpha-IZURONIDAZE, alpha-L-IZURONIDAZE, glucosamine-N-sulfatase, Alpha-N-acetyl glucosaminidase, acetyl-coenzyme-A:alpha-glucosamine-N-acetyltransferase, N-acetyl glucosamine-6-sulfatase, the beta-galactosidase, the beta-glucuronidase, N-acetyl galactosamine-6-sulfatase, and fusibility CD 4 are included. The array which carries out the code of these protein is known.

[0147]



The vocabulary and exogenism point out the array introduced into a cell by the approach indicated by this specification. An exogenous array can have the same array as the endogenous array which exists in a cell, or a different array.

[0148]

Preferably, a DNA array is a straight chain array.

A targeting array or arrays A targeting array or arrays are the arrays made into a target, for example, the DNA array which makes possible homologous recombination into the genome of the cell containing the gene made into a target. The vocabulary "a targeting array" and a "contiguity array" are used exchangeable in this specification. A targeting array is a DNA array which is homologous at the DNA array which generally usually exists in a cell genome which is obtained (that is, it is fully similar so that it may be the same into Cell DNA or a targeting array and Cell DNA may be able to pass through homologous recombination). For example, a targeting array may fully be homologous at the array which at least the imprint initiation section of :code or the non-code DNA, and the target gene has in the upstream, the array in the target gene, the array on the lower stream of a river of the imprint halt part of the target gene, or the array that exists in a genome through previous qualification. The targeting array or arrays to be used is collated and chosen as the part which is going to change the array made into the part or target which is going to insert the DNA array chosen.

[0149]

One or more targeting arrays are usable. Preferably, two targeting arrays adjoin the DNA array chosen. A targeting array in a gene or a coding sequence (The array of an exon and/or the intron) etc., The coding sequence of a gene is adjoined immediately (for example, 10, 5, 4, 3, and 2 or less than 1 nucleotide are included between a targeting array and the coding region of a gene). Or excluding the further nucleotide at all, for the coding sequence upstream of a gene, it is the upstream of the coding sequence of (an array or an endogenous promotor array of a non-[ upper ] coding region, etc. and genes), and a few may be left, and you may exist (array of an endogenous promotor's upstream etc.). or [ that a targeting array or arrays are known now ] -- or although a property decision is not made structurally, the thing which it maps by this contractor using a restriction enzyme, and can be determined [ the field of the array by which sequencing was carried out, and which is made into a target, and/or ] and for which an upstream field is included further is possible.

[0150]

It is possible to adjoin the coding sequence of an endogenous gene immediately and to insert the upstream or the DNA array which separates considerably and includes a control array using a targeting array. or -- or it is possible to permute the array which affects the structure or stability of RNA produced or protein further, to remove it by targeting, to add it, or to embellish it with another method. or [ for example, / improving the function, stability and/or translation possibility of an RNA molecule by embellishing the RNA stability element of an RNA molecule, a splice site, and/or a leader sequence ] -- or changing is possible. or [ increasing transportation of protein, secretion, or a functional characteristic ] -- or since it embellishes, protein arrays, such as a signal sequence, a pro peptide array, an active site, and/or a structural sequence, can also be changed An alteration, for example, correction, is possible for a protein array also by making into a target the part in the gene which carries out the code of the protein containing mutation, for example, point mutation, again.

[0151]

In one side face, a targeting array may be homologous at the part of Homo sapiens follicle-stimulating hormone beta (FSHbeta). FSH is a gonadotropic hormone which plays an indispensable role in maintenance and generating of oocyte and a sperm in normal reproduction physiology. In FSH, the latter is participating in the biological specificity of FSH including two subunits, and alpha and beta. the target site whose fixed targeting array is homologous -- the inside of the exon of an FSHbeta gene, and/or the intron -- the upstream of an FSHbeta coding region -- and -- immediately -- adjoining -- or the upstream of an FSHbeta coding region -- and a few may be left and you may exist. For example, the first thing (all targeting arrays when [ or ] there is only one targeting array into a structure) of two targeting arrays may be the genome field origin of the upstream of an FSHbeta coding sequence. for example, this targeting



array -- the part of the array number 1 -- for example, -- about -7,454 to -1,417 (array number 2) -- or -- about -- the nucleotide of the array origin corresponding to -696 to -155 (array number 3) which 20, 30, 50, 100, or 1000 follow at least may also be included. The second thing of two targeting arrays is good also considering the genome field of the coding sequence upstream as a target, or good also considering the exon or the intron of a gene as a target (for example, the part of the array numbers 2 or 3 may also be included too). An array usable although FSHbeta is made into a target is further indicated by the United States patent application 09th by which the contents are completely used for this specification / No. 305,639.

[0152]

A targeting array may be homologous at the part of Homo sapiens interferon alpha 2 (IFNalpha2). Interferon alpha constitutes the complicated gene family which contains in the short arm of a chromosome 9 the gene of 14 which forms a cluster. Neither containing IFNalpha2 gene of these genes has the intron. Interferon alpha is produced by many of other cell strains with a macrophage, a T cell, and a B cell. It has been shown that interferon alpha is effective in having the remarkable antiviral effectiveness and treating infection by papillomavirus, B mold and a hepatitis C virus, the vaccinia, the herpes simplex virus, the band-like varicella herpesvirus (herpes zoster varicellous virus), and rhinovirus.

[0153]

the target site whose fixed targeting array is homologous -- the inside of the coding region of IFNalpha2 gene -- the upstream of a coding region -- and -- immediately -- adjoining -- or the upstream of a coding region -- and a few may be left and you may exist. For example, the first thing (all targeting arrays when [ or ] there is only one targeting array into a structure) of two targeting arrays may be the genome field origin of the upstream of IFNalpha2 coding sequence. For example, this targeting array may also contain the part (for example, nucleotide which 20, 50, 100, or 1000 follow at least) of the array number 4 corresponding to -511 from the nucleotide -4074 of IFNalpha2 gene. The second thing of two targeting arrays is good also considering the genome field of the upstream of the coding sequence itself as an indicator. For example, the second targeting array may also contain the same exogenous coding region in the number codon of the beginning of IFNalpha2 coding sequence at the 3' edge. On the occasion of homologous recombination, an exogenous coding region is rearranged with the part made into the target of endogenous IFNalpha2 coding sequence. An array usable although IFNalpha2 are made into a target is further indicated by the United States patent application 09th by which the contents are completely used for this specification / No. 305,638.

[0154]

In another side face, a targeting array may be homologous at the part of a Homo sapiens granulocyte colony-stimulating factor (GCSF). GCSF is cytokine which stimulates the growth and differentiation of a hemopoietic precursor cell which are connected with neutrophil leucocyte / granulocyte cell lineage. GCSF is daily used in relation to prevention of the neutropenia which a chemotherapy guides, and a bone marrow transplantation. A chronic idiopathy and a congenital neutropenia failure also show an improvement after GCSF injection. the target site whose fixed targeting array is homologous -- the inside of the exon of a GCSF gene, and/or the intron -- the upstream of a GCSF coding region -- and -- immediately -- adjoining -- or the upstream of a GCSF coding region -- and a few may be left and you may exist.

[0155]

For example, the first thing (all targeting arrays when [ or ] there is only one targeting array into a structure) of two targeting arrays in a structure may be the genome field origin of the upstream of a GCSF coding sequence. For example, this targeting array may also contain the part (for example, nucleotide of the array origin corresponding to -364 (array number 6) from grade -6,578 which 20, 50, 100, or 1000 follow at least) of the array number 5 corresponding to 101 from the nucleotide -6,578 of a Homo sapiens GCSF gene. The second thing of two targeting arrays in a structure is good also considering the genome field of the coding sequence upstream as a target, or good also considering the exon or the intron of a gene as a target (for example, the part of the array number 6 may also be included



too). An array usable although GCSF is made into a target is further indicated by the United States patent application 09th by which the contents are completely used for this specification / No. 305,384.



[0156]

Control array A DNA array can include a control array. A control array may also include one or more promoters (a configuration or inductive promoter), an enhancer, UAS, a scaffold adhesion field or a matrix attachment site, an electronegative control member, a transcription factor binding site, or the combination of these arrays.

[0157]

It is possible for an inductive promoter to be included, for example, although a cell does not discover this product, guiding so that it may be discovered is possible, so that a cell may discover a product and it can be guided, in case it is produced, or in case a control array is introduced into an individual. An inductive promoter can be included by the control array so that a product may be discovered on the occasion of installation of a control array. A control array may be cellularity or a virus array. Although not necessarily limited to such a control array, what controls the early stages of SV40 or a late gene, an adenovirus main late gene, a mouse metallothionein-I gene, an elongation factor-1alpha gene, a cytomegalovirus gene, a collagen gene, an actin gene, an immunoglobulin gene, gamma actin gene, transcriptional activator YY1 gene, a fibronectin gene, or HMG-CoA reductase gene expression is contained. A control array can include further transcription factor binding sites, such as a TATA box, a CCAAT box, AP1 and Sp1, or a NF-kappa B binding site.

[0158]

Further DNA array element One or more exons can still be included by the DNA array. An exon is a DNA array which is copied to RNA and exists in a mature mRNA molecule. An exon carries out the code of the one or more amino acid, and DNA which carries out the code of/or the amino acid partially (namely, 1 or two bases of a codon) can be included. Or an exon contains DNA corresponding to a non-coding region, for example, a 5' non-coding region. When an exogenous exon or exons carry out the code of the parts of one or more amino acid and/or amino acid, a DNA array can be designed so that it may be the coding region and in frame of a gene by which a reading frame is made the second exon or target on the occasion of an imprint and splicing. In this specification, when the coding sequence of the first exon and the second exon unites in frame, it is the method to which a reading frame with the suitable part of mRNA originating in the second exon is not changed, and means connecting with both nucleotides.

[0159]

If the first exon of the gene made into a target includes the array ATG which starts a translation, an exogenous exon contains ATG preferably. Furthermore, the exogenous exon containing ATG may also contain one or more nucleotides further so that the coding region which mRNA containing the exon which the gene made into the second exon and target follows produced may be in frame. The gene which carries out the code of human erythropoietin, a human growth hormone, Homo sapiens colony stimulating factor-granulocyte / macrophage (hGM-CSF), and the Homo sapiens colony stimulating factor-granulocyte (hG-CSF) is contained in the example of the gene made into such a target in which the first exon contains ATG.

[0160]

A splice-donor part is an array which directs splicing from one exon to another exon. typical -- the first exon -- 5' of the second exon -- it is -- and 3' of the first exon -- the splice-donor part which overlaps the first exon in a side and adjoins -- 5' of the second exon -- the splice-acceptor part which adjoins the second exon in a side is recognized. Splice-donor part: It is AG (A/C) for which GU in the fourth and the fifth place is needed. It may have the characteristic consensus sequence indicated to be GURAGU (for R to show a purine nucleotide) (Jackson(1991) Nucleic Acids Res. 19:3715-3798). The first three bases of a splice-donor consensus part are three bases of the last of an exon. The capacity to attain a suitable reaction can define a splice-donor part functionally within a mRNA splicing path.

[0161]

A non-pair splice-donor part exists in the array made into a target, and is a splice [ which the splice-acceptor part arranged in a DNA array at 3' of a non-pair splice-donor part does not accompany ]-donor part. A non-pair splice-donor part can produce splicing to an endogenous splice-acceptor part.



[0162]

A splice-acceptor part is an array which directs splicing from one exon to another exon like a splice-donor part. Splicing equipment acts combining a splice-donor part and attains removal of the intron using a splice-acceptor part. As for a splice-acceptor part, Y may show one of pyrimidines :YYYYYYYYYNYAG and here, and it may have the characteristic array shown by supposing that N shows one of nucleotides (Jackson(1991) Nucleic Acids Res. 19:3715-3798).

[0163]

The intron is between two exons and is defined as an array of one or more nucleotides removed from a precursor RNA molecule by splicing on the occasion of formation of a mRNA molecule.

[0164]

Since a control array starts a translation, it can be connected with an ATG initiation codon. It is possible to connect a CAP part with a control array and an ATG initiation codon by request (about the specific mRNA initiation section used for this field in relation to regulatory region). Or in relation to a control array, the CAP part used for this array is not included in a target sequence, and imprint equipment offers a new CAP part. A CAP part can usually be found out to about 25 nucleotides 3' of a TATA box. Since a splice-donor part is the second exon and in frame of a gene which adjoin ATG immediately, for example, are made into a target, it can be arranged in the part by which existence of one or more nucleotides is not needed for an exogenous exon. DNA which carries out the code of the part of one or more amino acid which is the coding sequence and in frame of a gene which are made into a target, or amino acid may adjoin 3' side of ATG immediately, and may be arranged. As such a thing, a splice-donor part may adjoin 3' side of Code DNA immediately, and may be arranged.

[0165]

The code part (for example, exon 1 of a DNA array) of a DNA array may carry out the code of the parts of the one or more same amino acid as the thing of endogenous protein, and/or amino acid. For example, a code DNA array may correspond to the first exon of the target gene. Or for example, when the amino acid of the first exon of target protein is not important for the activity or activity of this protein, Code DNA may carry out the code of the part of one or more different amino acid from the first exon of target protein, or amino acid. For example, when building the fusant to an endogenous human erythropoietin (EPO) gene, the array which carries out the code of the first exon of a human growth hormone (hGH) is usable. In this example, the fusion to the EPO exon 2 of the hGH exon 1 produces formation of the functioning hybrid transit peptide. However, any exons of the Homo sapiens or the nonhuman origin which does not bar the function of hybrid transit peptide have the usable amino acid by which a code is carried out.

[0166]

When desirable products are endogenous protein and fusion protein of the coding sequence under DNA array, the exogenous code DNA incorporated by the cell may also contain DNA which carries out the code of one or more exons or array of cDNA corresponding to the translation or transcript which it is going to unite with the product of the gene made into an endogenous target. Therefore, it is possible to prepare the chimera or multifunctional protein which makes one polypeptide join ligand or an acceptor joint property that two or more protein is structural and like an enzyme using targeting. or [ for example, / that an exogenous DNA array offers mooring to the film of the protein made into a target, or cytokine secretion ] -- or the code of the transit peptide and the leader sequence which improve, an enzyme field, a film penetration domain field, a cofactor joint field, or other functional areas may be carried out. Usually, although not secreted, it is made to unite with signal protein and dopa decarboxylase, imprint control protein, and tyrosine hydroxylase are contained in the example of the protein which can offer secretion.

[0167]

It exists naturally or a DNA array can be obtained from the source of supply which can be produced using a genetic manipulation technique or a synthesis method.

Target sequence When transfection is carried out to cells, such as the founder, secondary, or an immortalization cell, a DNA array can adjust the activity of a desirable product, for example, protein, or



RNA, or the manifestation of a functional division. A DNA array may carry out the code of the desirable product again. A product may be hormone, cytokine, an antigen, an antibody, an enzyme, a coagulation factor, transporter protein, an acceptor, control protein, structural protein, a transcription factor, an antisense RNA, or a ribozyme. Furthermore, a product may be the protein or the nucleic acid (namely, fusion protein or a nucleic acid) which is not produced naturally.

[0168]

To such a product, erythropoietin, calcitonin, a growth hormone, An insulin, in SURINO tropine, an insulinlike growth factor, parathyroid hormone, Interferon beta and interferon gamma, a nerve growth divisor class, FSHbeta, TGF-beta, a tumor necrosis factor, glucagon, the bone growth factor -2, the bone growth factor -7, TSH-beta, Interleukin 1, interleukin 2, interleukin 3, interleukin 6, interleukin 11, interleukin 12, CSF-granulocyte, A CSF-macrophage, CSF-granulocyte / macrophage, and immunoglobulins Catalyst nature antibodies, protein kinase C, a glucocerebrosidase, superoxide dismutase, A tissue plasminogen activator, urokinase, Antithrombin III, DNase, the alpha-galactosidase, tyrosine hydroxylase, a blood coagulation factor V A blood coagulation factor VII, a blood coagulation factor VIII, a blood coagulation factor IX, a blood coagulation factor X A blood coagulation factor XIII, apolipoprotein E, or apolipoprotein A-I, Globins, a low-density lipoprotein acceptor, IL-2 acceptor, and IL-2 antagonists An alpha 1-anti trypsin, immunoreaction qualification agents, beta-GURUKO ceramidase, alpha-IZURONIDAZE, alpha-L-IZURONIDAZE, glucosamine-N-sulfatase, Alpha-N-acetyl glucosaminidase, acetyl-coenzyme-A:alpha-guru KOSAMIDO-N-acetyltransferase, N-acetyl glucosamine-6-sulfatase, the beta-galactosidase, the beta-glucuronidase, N-acetyl galactosamine-6-sulfatase, and fusibility CD 4 are included.

[0169]

A selectable marker and magnification Identification of a targeting event can be made easy by use of one or more selectable marker genes. These markers may exist in a construction lifter which may be contained in a DNA array or is different. a selectable marker -- two category:positivities -- selectable and shade sexual selection -- being possible (if it putting in another way marker for electropositive selection or one of shade sexual selection) -- it can divide. In electropositive selection, the cell which discovers an electropositive selectable marker a selection agent (neo and the xanthin-guanine phosphoribosyltransferase (gpt) --) dhfr, adenine deaminase (ada), puromycin (pac), Hygromycin (hyg), the carbamoylphosphate synthase, aspartate transcarbamylase, And CAD which carries out the code of the dihydroorotase glutamine synthetase (GS), It is possible to survive the processing in Histidine D (hisD) etc. in multiple drug resistance 1 (mdr1) list, and a targeting structure enables selection of the cell included in the host cell genome. In shade sexual selection, the cell which discovers a shade sexual-selection possible marker is destroyed under existence of a selection agent. Although a shade sexual-selection possible marker is connected with an exogenous DNA array, identification of a targeting event can be made easy by use of one or more marker genes which show the property of shade sexual selection which is arranged so that a right homologous recombination event with the array in a host cell genome may not produce the stable nest of a shade sexual-selection possible marker so that a targeting array may be adjoined (Mansour et al. (1988) Nature 336:348 -352). The antisense RNA or ribozyme of mRNA which carries out the code of the indispensable gene to a useful marker at a herpes-simplex-virus thymidine kinase (TK) gene, a bacteria gpt gene, a diphtheria toxin, and cell survival is contained in this purpose.

[0170]

You may also incorporate various selectable markers into the founder, secondary, or an immortalization cell. For example, the selectable marker which gives selectable phenotypes, such as a manifestation of drug tolerance, auxotrophy, the resistance over a cell trauma agent, or surface protein, may be used. neo, gpt, dhfr, ada, pac, hyg, CAD, GS and mdr1, and hisD are contained in an usable selectable marker gene. The given selectable phenotype identifies a recipient cell and makes it possible to isolate.

[0171]

The gene (for example, ada, GS, dhfr, and a multifunctional CAD gene) which carries out the code of the selectable marker has the added property which enables selection of a cell including the selectable



marker and the adjoining genome array of a copy which increased. This description offers a device for the copy which increased to make the copy number of a desirable contiguity or a desirable connection gene increase intentionally. Other arrays which draw the mutant and the copy which increased of these arrays that show the improved selection property are usable.

[0172]

The sequence and the number of components under DNA array may be various. For example, sequence: It is [ -- It is a splice-donor part. / -- You may be the second targeting array, or it sets to an alternative, and is the first targeting array / -- It is a control array. / -- It is an exon. / -- It is a splice-donor part. / -- It is DNA which carries out the code of the selectable marker. / -- You may be the second targeting array. ] the first targeting array. -- It is a selectable marker. -- It is a control array. -- It is an exon. The subset of the cell by which the cell which builds a structure into stability will survive processing by the selection agent, and transfection was carried out to; stability is a homologous recombination cell. A homologous recombination cell can be identified with various techniques including PCR, Southern hybridization, and phenotype screening. Sequence of a structure: It is [ -- It is a splice-donor part. / -- It is the intron. / -- It is a splice-acceptor part. / -- You may be the second targeting array. ] the first targeting array. -- It is a selectable marker. -- It is a control array. -- It is an exon.

[0173]

Or the sequence of the component under DNA array is [ -- It is a splice-donor part. / -- It is the second targeting array. / -- You may be the selectable marker 2 or it is the first targeting array. / -- It is a control array. / -- It is an exon. / -- It is a splice-donor part. / -- It is the selectable marker 1. / -- It is the second targeting array. / -- You may be the selectable marker 2. ] the targeting array of :first. -- It is the selectable marker 1. -- It is a control array. -- It is an exon. By this arrangement, the selectable marker 2 may show the property of shade sexual selection. That is, the gene product of the selectable marker 2 is selectable on the contrary by the growth under suitable culture-medium formula containing \*\* (typically drugs or a metabolite analog) which kills the cell which has discovered the selectable marker 2. The recombination between the targeting array which adjoins the selectable marker 1, and the homologous array in a host cell genome produces the nest made into the target of the selectable marker 1, and, on the other hand, the selectable marker 2 is not incorporated. Although it rearranges and transfection of the event is carried out to stability with the selectable marker 1, such a cell by which transfection is not carried out to stability with the selectable marker 2 is generated, and such a cell is selectable by the growth in the culture medium containing the selection agent chosen in opposition to the selection agent and the selectable marker 2 which are chosen in support of the selectable marker 1.

[0174]

A DNA array can also include a positivity [ which enables selection of the cell which includes the copy which the marker increased again ] selectable marker. The copy which such a marker increased produces coincidence magnification of the adjoining DNA array. For example, sequence of a component: It is [ -- It is an exon. / -- It is a splice-donor part. / -- You may be the second targeting DNA array. ] the first targeting array. -- It is the electropositive selectable marker to which a copy number is made to increase. -- It is the second selectable marker (based on a request). -- It is a control array. As for the gene activated, it is possible for a cell including the copy which the selectable marker gene increased to make it increase further by inclusion of a selectable marker gene in which it has a selectable property by cultivating a cell under existence of a suitable selectable agent. It will be increased to the selectable marker gene and the tandem by the endogenous gene activated. It is possible for a multi-copy \*\*\*\* cell to produce protein with a very desirable high level for the endogenous gene activated, and it is in. It is useful to vitro protein production and gene therapy.

[0175]

The marker gene of selectable and others does not immediately need to adjoin mutually.

DNA -- an array / a homologous recombination improver / non-homologous edge connection inhibitor complex The homologous recombination between the targets DNA DNA, for example, the chromosome in a cell, chosen [ which are chosen and are double-stranded-DNA-arranged ] can promote by approaching the part made into a DNA array and a target very much enough, and offering \*\* which



increases homologous recombination, for example, Rad52 protein, and \*\* which checks non-homologous edge connection, for example, Ku deactivator, (for example, anti-Ku antibody). The concentration of \*\* from which "approaching very much enough" prevents a homologous recombination improver and/or non-homologous edge connection in this specification points out installation of \*\* which checks a sufficient homologous recombination improver or non-homologous edge connection to offer the homologous recombination between the alteration of a part made into a target, for example, a DNA array, and a target sequence at a higher rate, or both. It is possible to offer the installation which a DNA array, a homologous recombination improver, and \*\* that checks non-homologous edge connection approached very much mutually using some approaches. A homologous recombination part is made to carry out localization of the activity of compounds, such as Rad52 and Ku inactivation molecule, for example, anti-Ku antibody etc., by approaching Target DNA very much and medicating him with these compounds of each other [ and ]. For example, partial inhibition of Ku activity may be more desirable than all cell inhibition of Ku activity.

[0176]

Contiguity of a DNA array, a homologous recombination improver, and \*\* that checks non-homologous edge connection is maintainable by introducing these elements as some complex. For example, DNA-protein complex is usable. The core of DNA-protein complex may consist of double-stranded-DNA arrays which it is going to introduce into the part to which Target DNA is chosen. Even if a DNA array top, for example, all the arrays or edge of a DNA array, has few single strand protrusion edges of for example, a DNA array, on the part, a homologous recombination improver, for example, Rad52 protein, or its fragment adheres, for example, coating may be carried out to it. Ku deactivators, such as \*\* which checks the non-homologous edge connection which is carrying out covalent bond to a DNA array or one of homologous recombination improvers, for example, anti-Ku antibody etc., can be further included by DNA-protein complex. The noncovalent bond of the \*\* which checks non-homologous edge connection may be carried out to the DNA array or the homologous recombination improver again.

[0177]

By providing into liposome or a vesicle, mutually, it approaches very much and a compound can also maintain \*\* which checks a DNA array, a homologous recombination improver, and non-homologous edge connection again. For example, the liposome suspended solid is usable also as a carrier which can be permitted pharmacologically [ these elements ] again. A liposome suspended solid can be prepared according to the approach learned by this contractor so that it may be indicated by U.S. Pat. No. 4,522,811.

[0178]

\*\* which check a DNA array, a homologous recombination improver, and non-homologous edge connection may be some mixed solutions which can carry out a microinjection to a cell again, or all compound three [ these ] exist in a cell at coincidence -- as -- these compound each -- other things -- then, you may introduce quickly. Acceptor agency conveyance, electroporation, and calcium phosphate precipitate are included in other approaches of introducing these one or more compounds.

[0179]

Cell The founder and the secondary cell which are going to carry out transfection can be obtained from various organizations, and are under culture, and contain the cell strain in which maintenance and growth are possible. For example, the element (for example, a lymphocyte, a bone marrow cell) with which fibroblast, a keratinization cell, an epithelial cell (for example, a mammary gland epithelial cell, an intestinal epithelium cell), an endothelial cell, a neuroglia, a nerve cell, and blood were formed, muscle cells, and these somatic cell kinds of precursor is contained in the founder and the secondary cell in which transfection is possible. Primary cell culture is preferably obtained from the individual medicated with the transfection founder or a secondary cell (namely, self-cell). However, primary cell culture may be obtained from the donor (except a recipient) of the same kind (namely, allogeneic cell) or another kind (namely, different-species cell) (for example, a mouse, a rat, a rabbit, a cat, a dog, Buta, a cow, a little bird, a sheep, a goat, a horse).

[0180]



Transfection is carried out in the vertebrate and exogenism DNA array to which the founder or the secondary cell of the mammalian origin carries out the code of an exogenous DNA array, for example, the therapy protein, preferably, and it is in vitro and in vivo and it is possible to produce the extended period and the therapy protein by which a code is carried out possible [ stability and reappearance ]. Furthermore, the transfection founder and a secondary cell are suitable level physiologically, and are in vivo, and can discover the product by which a code is carried out, and collect cells after transplantation, and it is made to increase on the occasion of re-culture, and it is possible for the property before the transplantation to be shown.

[0181]

Or transfection is possible in the vertebrate and exogenism DNA array in which the founder or the secondary cell of the mammalian origin includes a control array preferably. One or more :promoters and enhancers, UAS, a scaffold adhesion field, or an imprint binding site is included in the example of such a control array. A targeting event can put under the accommodation the endogenous gene which produces insertion of the control array of a DNA array and is made into a target (for example, thing for which a promoter, an enhancer, or both are inserted in the upstream of an endogenous gene or regulatory region). It is possible for a targeting event to produce deletion of an endogenous control array, such as deletion of the organization specific negative control array of a gene, in coincidence by request. A targeting event can permute the existing control array and can be permuted by the enhancer which shows the enhancer which has cell species-specific which is more larger than a natural existence element, or is different from it in; for example, an organization specific enhancer, different control from a corresponding non-transfection cell, or an induction pattern. In relation to this, deletion of the natural existence array is carried out, and a new array is added. Or although an endogenous control array is not removed or permuted, it is destroyed by the targeting event that it is because targeting of the exogenous array is carried out into an endogenous control member etc., or it is made into inability. Installation of the control array by homologous recombination may produce the founder or the secondary cell which discovers the therapy protein which is not usually discovered. or [ furthermore, / that the installation to which targeting of the control array is carried out creates therapy protein ] -- or normal, although contained -- or [ the cell which is small quantity (physiologically normal amount of under lower level) more, or is a deficit mold, and / increasing the content or production, although therapy protein is physiologically created on normal level ] -- or it is usable into the cell which it is going to increase.

[0182]

The transfection founder or a secondary cell may give selectable phenotype, and may also include the DNA array which makes the identification and isolation easy and which carries out the code of the selectable marker again. they are abnormalities through use of the ensemble of the transfection founder or a secondary cell at the approach and list which produce the clone sexual cell stock of the transfection founder who discovers a DNA array to stability, the approach of producing a secondary cell, and such a transfection cell and a heterogeneity cell strain, clone nature, and a heterogeneity cell strain -- again -- \*\* -- or [ treating the condition which is not desirable ] -- or the approach of preventing is a part of this invention.

[0183]

Production of the transfection of the founder or a secondary cell, homologous recombination and clone nature, or a heterogeneity cell strain Vertebrate tissue can be obtained by the standard approaches of obtaining the organization source of supply of a punch biopsy or the target primary-cell-culture kind, such as other methods of operation. For example, a punch biopsy is used for obtaining the skin as a source of supply of fibroblast or a keratinization cell. The mixture of primary cell culture is obtained from an organization using known approaches, such as enzyme digestion or explantation. When enzyme digestion is used, enzymes, such as collagenase, hyaluronidase, dispase, pronase, a trypsin, elastase, and a chymotrypsin, are usable.

[0184]

Before carrying out direct transfection of the produced primary-cell-culture mixture or performing transfection, it cultivates first, and it may remove from a culture plate and you may re-suspend. Together



with DNA containing DNA which carries out the code of the selectable marker by request which it is going to introduce into a genome, primary cell culture or a secondary cell is processed in order to attain transfection. Furthermore, it is independent of Rad52 protein or its fragment and Ku inactivation molecule, for example, anti-Ku antibody, and the founder or a secondary cell is doubled as some complex.

[0185]

The transfection founder or a secondary cell can be created by electroporation. Electroporation is performed with suitable potential and electric capacity (and corresponding time constant), and produces penetration of the DNA structure (kind) to the founder or a secondary cell. Electroporation can be performed over the potential (for example, 50 to 2000 volts) and the corresponding electric capacity of the large range. Usually, the total DNA of about 0.1 to 500microg is used.

[0186]

Preferably, transfection of the founder or the secondary cell is carried out using a microinjection. Or it is possible to carry out transfection of the cell using known approaches, such as calcium phosphate precipitate, qualification calcium phosphate precipitate and polybrene precipitate, liposome fusion, and acceptor agency gene conveyance. A stable transfection cell is isolated, it cultivates the bottom of a culture condition, and sufficient time, and subculture is carried out, and a stable transfection secondary cell is proliferated, and the clone sexual cell stock of a transfection secondary cell is produced. Or subculture of more transfection cells than 1 is cultivated and carried out, and production of a heterogeneous cell strain is produced.

[0187]

A cell is maintained under the conditions which make homologous recombination possible so that it may be known by the technical field concerned after transfection (Capecchi(1989) Science 244:1288-1292).

[0188]

The homologous recombination founder or a secondary cell can produce one of the clone sexual cell stocks or heterogeneity cell strains of sufficient size to be an effective dose and provide an individual with therapy protein through a sufficient number of doubling. Generally, the skin of 2 is obtained by the biopsy 0.1cm, and it is thought that this contains 100,000 cells, and a clone sexual cell stock is produced using a;1 \*\* cell, and the homologous recombination secondary cell of 100 million is produced through doubling of about 27. When it is going to produce a heterogeneity cell strain from the original homologous recombination ensemble of about 100,000 cells, only doubling of 10 is needed for producing the cell of 100 million.

[0189]

The number of the cells needed for homologous recombination clone nature or a heterogeneity cell strain is various. And use of a homologous recombination cell, the functional level of the exogenous DNA array in a cell although not necessarily limited, The various factors in a cell which include a patient's age, surface area, and a clinical condition in the functional level of the changed DNA array, the transplantation part (for example, the usable number of cells is restricted by the anatomical part of transplantation) of a homologous recombination cell, and a list are accepted. if these factors are put on a visual field, it has an isolated growth hormone deficiency, otherwise, in order to convey the therapy level of a human growth hormone to a 10kg healthy patient, the homologous recombination fibroblast of 500 million will be required from about 1 (the volume of these cells -- almost -- a patient's thumb -- very -- tip extent -- it is).

[0190]

It is possible to determine the effectiveness in which the approach indicated by this specification increases homologous recombination in a cell using some approaches. For example, since the non-conservative substitution in a cell, for example, a human cell, is detected, it is possible to design an experiment system. A permutation may be a permutation of C to T in the CGA codon of the exon 3 of the HPRT gene which is a part of XhoI part. This mutation generates the conclusion signal of TGA, and this produces the HPRT shade sexual expression mold by which score attachment is carried out with resistance in 6-thioguanine (6-TG). This mutation is accompanied by the corresponding deletion of a



XhoI part again. Briefly, a DNA array including the permutation of C to T can be introduced by carrying out a microinjection to Homo sapiens fibroblast as some complex containing \*\* which inactivates \*\* which increases homologous recombination, and Ku. It is made to increase before introducing a cell on the culture medium which cultivates a cell and contains 6-TG. Then, score attachment of the 6-TG resistance clone is carried out, and existence of a mutation DNA array is determined. Existence of a homologous recombination event is detectable using a HPRT specific probe by carrying out southern blot analysis of the XhoI digestive genomic DNA. A result can also be compared with the reference cell introduced into the bottom of the nonexistence of \*\* to which a mutation DNA array inactivates \*\* which increases homologous recombination, and Ku again.

[0191]

Transplantation of the clone sexual cell stock of a homologous recombination secondary cell, or a heterogeneity cell strain The homologous recombination cell produced as mentioned above can be introduced into the individual which is going to convey therapy protein using a known approach. Then, a clone sexual cell stock or a heterogeneity cell strain is introduced into an individual using various routes of administration and various parts (for example, the inside of the bottom of the renicapsule, hypodermically, a central nervous system (the inside of a subarachnoid cavity is included), a blood vessel, liver, and internal organs, intraperitoneal (the inside of a hawser is included), or muscle implantation graft) using a known approach. If transplanted to an individual, a homologous recombination cell will produce the therapy product by which a code is carried out to an exogenous synthetic DNA, or a homologous recombination cell will discover the therapy protein by which a code is carried out to an endogenous DNA array under exogenous control array accommodation. For example, the individual which is the bleeding disorder caused to blood by the deficit of the factor IX usually seen and which is diagnosed as hemophilia B is the candidate of gene therapy. ; which performs a pellicle skin biopsy to a patient -- this is the simple treatment which can be carried out based on an outpatient department. About 1 minute is mostly taken for this to remove by taking the piece of the skin of the magnitude of the head of a match from under an arm. Processing of the sample is carried out, and genetic manipulation is carried out so that a patient cell (fibroblast in this case) may be isolated and the lost factor IX may be produced. Based on a patient's age, weight, and a clinical condition, a number of cells needed are proliferated by large scale cultivation. All processes should require four to six weeks, and introduce into an individual a suitable number of cells by which genetic manipulation was again carried out to the end of this period by the outpatient department (for example, thing for which these are poured in and returned to the bottom of a patient's skin). A patient can produce the own factor IX here, and is not a hemophiliac any longer.

[0192]

It is possible to treat other abnormalities or diseases using the same approach. For example, low height can be treated by medicating an individual with a human growth hormone by transplanting the founder or the secondary cell which discovers a human growth hormone.

[0193]

Generally the cell to be used is a cell by which genetic manipulation was carried out to the patient unique target, as this example suggests. However, it is possible to obtain a cell from another individual of the same kind or a different kind. Use of such a cell may require use of the obstruction equipment which prevents administration of an immunosuppressant, the alteration of a histocompatibility antigen, or refusal of a transplantation cell.

[0194]

Probably with many diseases, this is a single time therapy, and many gene therapy measures are needed in other things.

The transfection founder or a secondary cell is independent, or can be prescribed for the patient combining the obstruction for checking the immunoreaction to the cell in a recipient test subject, or \*\*. For example, it is possible to medicate a test subject with an immunosuppressant, and to check the normal reaction in a test subject, or to interfere with this reaction. Preferably, an immunosuppressant is an immunosuppressive agent which checks T cell/or B cell activity in a test subject. The example



of such an immunosuppressive agent is commercially available (for example, Cyclosporin A is commercially available from Sandoz Corp. and New Jersey East Hannover).

[0195]

An immunosuppressant, for example, drugs, is sufficient dosage to attain the desirable therapy effectiveness (for example, inhibition of refusal of a cell), and a test subject can be medicated with them. The medication range of an immunosuppressant is known by the technical field concerned. Freed et al. [ for example, ] (1992) N. Engl. J. Med. 327:1549; -- Spencer et al. (1992) N. Engl. J. Med. 327:1541; -- Widner et al. (1992) N. Engl. J. Med. Please refer to 327:1556. A medication value may be various according to factors, such as a disease condition of an individual, age, sex, and weight.

[0196]

Another \*\* usable although T cell activity is checked in a test subject is an antibody, its fragment, or a derivative. in or [ making a T cell drained in vivo ] -- or the antibody which can be isolated is known by the technical field concerned. Polyclonal antiserum, for example, antilymphocyte serum, is usable. Or one or more monoclonal antibodies are usable. The monoclonal antibody combined with CD2, CD3, CD4, CD8 and CD40 on cell surface, and CD40 ligand is contained in a desirable T cell exhaustion antibody. Such an antibody is known by the technical field concerned, and is commercially available from for example, an American type culture collection. The desirable antibody combined with CD3 on a Homo sapiens T cell is OKT3 (ATCC CRL 8001).

[0197]

or [ whether a T cell is made drained within a recipient test subject, or that it is isolated ] -- or the antibody to check is the dosage which checks refusal of a cell on the occasion of transplantation, and can prescribe [ a suitable period and ] a medicine for the patient An antibody is administered intravenously in the carrier which can be permitted pharmacologically preferably, or a diluent (for example, physiological saline solution).

[0198]

In a recipient test subject, the option which interferes with the immunoreaction to a cell or checks this reaction is using an immunity obstruction. An "immunity obstruction" points out the equipment which serves as an obstruction between the cells which participate in the cell and immunoreaction in a test subject which were prescribed for the patient in this specification. For example, a cell can prescribe a medicine for the patient in implantable equipment. Although a semipermeable obstruction, i.e., a nutrient, and a product diffuse implantable equipment in an obstruction and out of an obstruction, it may also contain the cell contained inside the obstruction which bars penetration of a larger immune system component, for example, an antibody, or complement. Typically, implantable equipment contains the core which arranges a matrix, for example, a hydrogel, or a cell. By request, semipermeable coating may enclose gel. When arranged gel incore, the cell prescribed for the patient should be isolated from the cell of an immune system, and should be covered from a host's cell and cytotoxic antibody. Preferably, permselectivity coatings, such as PLL or PLO, are used. A recipient's immune system component advances and coating often has the stoma which bars destroying the cell in implantable equipment.

[0199]

Many approaches of wrapping a cell entirely are learned by the technical field concerned. For example, in order to wrap the cell for production entirely, the encapsulation and other encapsulation methods for obtaining semipermeable water-insoluble nature gel are indicated by U.S. Pat. No. 4,352,883 using water-soluble rubber (gum). Other usable implantable equipments are U.S. Pat. No. 5,084,350, U.S. Pat. No. 5,427,935, and WO released to 27 in July, 1995. It is indicated by 95/19743, U.S. Pat. No. 5,545,423, U.S. Pat. No. 4,409,331, U.S. Pat. No. 4,663,286, and the Europe patent No. 301,777.

[0200]

Use of the homologous recombination founder, a secondary cell, and a cell strain The homologous recombination founder and a secondary cell, or a cell strain has large applicability as the vehicle or conveyance system for therapy protein, such as Actuation DNA, in an enzyme, hormone, cytokine, an antigen, an antibody, a coagulation factor, an antisense RNA, control protein, imprint protein, an acceptor, structural protein, new (un-optimizing) protein and a nucleic-acid product, and a list. For



example, although not necessarily limited using the homologous recombination founder or a secondary cell Erythropoietin, calcitonin, a growth hormone, an insulin, in SURINO tropine, An insulinlike growth factor, parathyroid hormone, alpha2-interferon (IFNA2), A beta interferon, a gamma interferon, a nerve growth divisor class, FSHbeta, TGF-beta, a tumor necrosis factor, glucagon, the bone growth factor -2, the bone growth factor -7, TSH-beta, Interleukin 1, interleukin 2, interleukin 3, interleukin 6, interleukin 11, interleukin 12, CSF-granulocyte (GCSF), A CSF-macrophage, CSF-granulocyte / macrophage, and immunoglobulins Catalyst nature antibodies, proteinkinase C, a glucocerebrosidase, superoxide dismutase, A tissue plasminogen activator, urokinase, Antithrombin III, DNase, the alpha-galactosidase, tyrosine hydroxylase, a blood coagulation factor V A blood coagulation factor VII, a blood coagulation factor VIII, a blood coagulation factor IX, a blood coagulation factor X A blood coagulation factor XIII, apolipoprotein E, apolipoprotein A-I, Globins, a low-density lipoprotein acceptor, IL-2 acceptor, and IL-2 antagonists An alpha-1-anti trypsin, immunoreaction qualification agents, beta-GURUKO ceramidase, alpha-IZURONIDAZE, alpha-L-IZURONIDAZE, glucosamine-N-sulfatase, Alpha-N-acetyl glucosaminidase, acetyl-coenzyme-A:alpha-glucosamine-N-acetyltransferase, It is possible to supply N-acetyl glucosamine-6-sulfatase, the beta-galactosidase, the beta-glucuronidase, N-acetyl galactosamine-6-sulfatase, and therapy protein including fusibility CD 4.

[0201]

As for all the patents and bibliographies that are quoted by this specification, the whole is used for this specification. Other modes are within the limits of a claim.

[Layout Table]



## SEQUENCE LISTING

&lt;110&gt; Evguenii Ivanov

&lt;120&gt; METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION

&lt;130&gt; 10278/016001

&lt;160&gt; 9

&lt;170&gt; FastSEQ for Windows Version 3.0

&lt;210&gt; 1

&lt;211&gt; 7622

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 1

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ggatccgaga  | acatagaagg  | agcaggtaat  | ttatcaaggg  | atgaacacgg  | gtgcttaatt  | 60   |
| tootattttg  | aggocaggca  | tggtggctca  | caactgtaat  | occaacactt  | taggaagcoa  | 120  |
| aggtgggtgg  | attgottgag  | totaggattt  | tgagacacgg  | otggccaaac  | tgccgaatco  | 180  |
| otgtotctac  | taaaaatact  | aaaattaacc  | agtcattggt  | gtgggtgtgc  | tttagtccca  | 240  |
| gctactctgg  | tggttgaggc  | acaagaatca  | cttggaacct  | ggaggcagag  | gttgacgtga  | 300  |
| gctgagactg  | tgcoacttoa  | ctccagcctg  | ggtgacagag  | taagattctg  | totcaaaaaa  | 360  |
| tatgtatata  | tacacacata  | taatagatac  | ataaacatat  | atacatatat  | aatatatata  | 420  |
| tatatatatt  | tatatataata | tataaacata  | tataaatata  | tatatatata  | tatatatata  | 480  |
| tataaaac    | aaacataaag  | gaataatttt  | gggggaaaaa  | cttcaataat  | gaagaaacaa  | 540  |
| cataggtgtg  | tgagtatatg  | cacagaaatt  | caagagatct  | toacagcaat  | gaagacattg  | 600  |
| gtttacocaga | attcaaaaaa  | gaagtoagct  | gtgcatttaa  | agtagaatgt  | gatgagtgtt  | 660  |
| accactgagg  | taggaactgg  | gaactaagg   | agcgttaagc  | agaaagtgtc  | gaactgagag  | 720  |
| ttgggcattg  | gaggtgtgtg  | aaggcagggt  | aagtgaatgt  | ctoctagaag  | ctacotttaa  | 780  |
| atggagtttt  | gaagtacttg  | tagggagtagc | ttagggtgaa  | agaagaggag  | aaacatgtat  | 840  |
| caggcagagg  | gaotagaacc  | ttattacott  | caaagaagaa  | gcaaaaagaa  | tacattgtgc  | 900  |
| tttgagggtg  | tgagggtgac  | tttaagccaa  | tataaggtgaa | tttgacatag  | gaattcccta  | 960  |
| aatattgttc  | ggtcaattgt  | taaatattga  | gtgatataac  | actgtattaa  | agcccaagag  | 1020 |
| ttgcttttat  | atagaaagaa  | gaaaaaagcc  | caagagagtt  | ttatttctag  | agggaaatatt | 1080 |
| ttctagaaat  | aaagggaagg  | gtatcagoca  | gtttctagtc  | aggaaaacag  | aaatcacacc  | 1140 |
| tgatatgoaa  | aatagaggaa  | aatcaggga   | ttcatttaac  | cagagatttg  | gttgctcaag  | 1200 |
| tattagattg  | ctgaaaagoc  | agacaggga   | tatgaggcaa  | toagagataa  | gtattagtga  | 1260 |
| caagctccat  | ttatgtgcag  | gattggagg   | acataggtgg  | gggtcccaag  | agccagaaag  | 1320 |
| tgagaccacc  | tagcagaagg  | tcaaaaccac  | gctggggttt  | ctcacaaaaa  | gotgggacca  | 1380 |
| ccaggaggag  | ctgtccaatg  | ggatctggag  | ccaggagat   | catgagtgca  | ctaccaggaa  | 1440 |
| gggaagcaga  | atgtaaaagg  | tagagagaaa  | tactccaaot  | gcttcoottgc | attcaotttc  | 1500 |
| caatctccat  | tcacaaaagg  | aaaaaactgc  | taatacagca  | gagtgggaaa  | agcagcctgc  | 1560 |
| caaggtcoot  | totcccacaa  | aacagagcac  | aaaaccaaag  | aaaaacaaag  | aatgcatttg  | 1620 |
| atagcaaaac  | ggctatggac  | caacccaaca  | taaaagaat   | gatgagtgat  | ttcttttttc  | 1680 |
| atttggttoa  | agaaaagtat  | ttcagtaaat  | attatgtaac  | agaaattcta  | tttattttgg  | 1740 |
| ggaattcaaa  | ggtgaataaa  | aaagaactot  | aaatttttat  | caataaata   | tttcaaaaac  | 1800 |
| ctcaatgaga  | gtaatggcat  | taactagcaa  | atatgcta    | gagatgagct  | agccataaga  | 1860 |
| ggottagaat  | tgagagaaag  | gtctgggggc  | ctcttgacag  | gccaattcca  | gagotgtttg  | 1920 |
| tggaatctc   | tgacottaat  | gcagggtgaa  | atataaatat  | gggcatttag  | aatagtggcc  | 1980 |
| caacttttg   | atgattttot  | tottggtgto  | totccaatta  | atgggattga  | tgagaactgt  | 2040 |
| agaccactga  | ggtcaaccat  | gctcaatgaa  | tagtccootg  | gctttggagt  | caaatgacc   | 2100 |
| tgaaatagaa  | ccccagottt  | gctacttaca  | ggttgcattt  | atcctcagtt  | ttctcatott  | 2160 |
| tcaaaagaaga | acagtaactt  | otttaaaagg  | ttattgttag  | otgggtgcag  | tggtccacgc  | 2220 |
| ctgtaactgc  | agcacttttg  | gagggcgagg  | ctagtggatc  | acttgaggcc  | aggagttgga  | 2280 |
| aactagootg  | gocaaacatg  | tgaaactctg  | tototacaaa  | agaaatttta  | aaaaattttg  | 2340 |
| ctgggtgtgg  | tgccacacac  | ctggaattcc  | agotacctgg  | gagggccagg  | catgagcatc  | 2400 |
| aottgagtot  | ggaaagcaga  | gggttgaggt  | gagocaaagt  | tgtaacactg  | tactcaagcc  | 2460 |
| tggttgacac  | agtggagacot | tgtctaaaaa  | aaaaaaagg   | tattgtgtta  | ttgtaaatat  | 2520 |
| tgtatatgaa  | ottctatttta | aatgttttag  | ttaaatgoot  | gtgtaattgt  | ocaaatgtgt  | 2580 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ottotagotc  | actgcacaga  | aaaaactgat  | tcactgaaat  | oatggaattg  | cagcaaagaa  | 2640 |
| caaatotaat  | taatgtaggt  | caaacgggag  | gactggagtt  | attattcaaa  | toagtctccc  | 2700 |
| tgaaaaotca  | gaggctaggg  | ttttatggat  | aatttgggtg  | gcaggggaot  | aggggaatggg | 2760 |
| tgctgctgat  | tggttgggga  | atgaaatagt  | aagattgtgg  | aaaactgtcc  | tccttcattg  | 2820 |
| agtctgcttc  | cgggtgtagg  | ccacacgacc  | agttgagtoa  | tgaagcatgc  | gtccaaagtgg | 2880 |
| agtcagtttg  | ttgccagaaat | gcaaaaagcct | gaaaaatgtc  | tczaatgatc  | aaotgtaggc  | 2940 |
| tcacacataa  | tgtattatct  | tataggagca  | attgggggaag | taacaaatct  | tgtgacctct  | 3000 |
| ggcaacataa  | ctcctgaact  | agtaagggat  | tataaaaaac  | atgcoctatat | cttatcagaa  | 3060 |
| ttcagggtccc | cccataatcc  | taatotcaca  | goatttcatt  | tgtttagaaa  | ggocattttc  | 3120 |
| agtccttgag  | caaggagggg  | gttagtttta  | ggataggact  | attatccttg  | ctcogttaaa  | 3180 |
| ctataaacta  | aattcctccc  | atggttagot  | tgcoctaac   | otaagaatga  | gtgagaacag  | 3240 |
| coagcctgtg  | aggotagagg  | caagatggag  | tcagccatgc  | tagatttato  | toactgtcat  | 3300 |
| aaotctttgca | aaggcagttt  | caoctgggac  | ataggaggta  | ctcaatgaaa  | aagaagctat  | 3360 |
| taattattaaa | attttaaaaa  | tgaatttaag  | gaaotaatc   | tatgtacata  | ttagtcatta  | 3420 |
| aaacaaagtg  | gttoattttac | attcacacaa  | ataaatcttg  | tgattataca  | taggtaatat  | 3480 |
| gaaaaacttt  | gttttcttto  | ataatacaag  | gtattagcaa  | tagatatagt  | aatgttagca  | 3540 |
| ttccttttgga | aaaaatgaaa  | agattttata  | ttttccaaaga | atcatttagta | ttttttattta | 3600 |
| atatacataa  | tataaaattt  | attcattcta  | taacttggaa  | atatgottgc  | ttaccaatta  | 3660 |
| ctgacagatt  | tcaaaatatt  | totataotca  | caatattcoat | ttacataaat  | attgatttgg  | 3720 |
| taacttaaat  | gtgtactgot  | atgctaagtt  | ttgtotttgt  | caaaotatt   | ttataaaato  | 3780 |
| ataatcctag  | atgaatccaa  | ottttggtaa  | cccaogtgc   | tgaaocctcg  | ctgttaaocg  | 3840 |
| gcaaatgttg  | gtaggtacgt  | attctatcct  | accacotctc  | tctacccacc  | agcatctgca  | 3900 |
| cccaacaccc  | ctccccaccc  | accattatct  | atacacaaca  | ccctcccaaa  | cccaacacga  | 3960 |
| totgcaacca  | cccaacogcc  | cacccaacac  | catgtacact  | caactacact  | cccaacacga  | 4020 |
| acccatctgca | cccatcactc  | ctcccatcc   | acaagcatct  | gcaocccaca  | caattcccta  | 4080 |
| ccaccagca   | totcactca   | cccaotctcc  | cccaacagcc  | atctgcaccc  | acaaocctc   | 4140 |
| ctcaccaccc  | agagtctgca  | tccatcacac  | ttgcccactc  | gotagcatct  | gcaccatcaa  | 4200 |
| gctctgocct  | ottgocctaat | acgggatgag  | ctctccatgg  | ttctgcctaa  | agacaatgct  | 4260 |
| tcactcctcc  | ttctataacc  | caattccttt  | taactotcca  | agtacaotcc  | agaaotctcc  | 4320 |
| tctcctctctg | ataccaactt  | tttccacttt  | actcactcat  | tcctatcacc  | atacaaaagt  | 4380 |
| gtttattttct | cccatottaa  | agttaaaaat  | caaaagaaaa  | ttgtctgggg  | ccaggcacgg  | 4440 |
| tggtctaacgc | ctgtaatccc  | aacactttgg  | gaggccaaag  | aggggtggat  | gaotaaaggt  | 4500 |
| taggagttca  | tgcatacagc  | ggccaacatg  | gtgaaaccca  | totctaotaa  | aaatacaaaa  | 4560 |
| attagccagg  | catggtggca  | catgcctgta  | gtctcaggta  | ottggggagg  | tgaggccaga  | 4620 |
| gaatggcttg  | aaocogggag  | gcagaggttg  | caatgagcog  | agattgtgco  | cttgcactcc  | 4680 |
| agcctgggtg  | acagagtgag  | actccatctc  | aaaaataaaa  | aatataaata  | aaacaaaaga  | 4740 |
| aagtattttt  | tacccaacat  | ccacattaac  | caataacca   | tttctttatt  | gatctttgta  | 4800 |
| aaaaaaagot  | cttggaaaaa  | ttgtctatat  | taactatgac  | ttatctctcc  | caaatcactt  | 4860 |
| aaacacatac  | caatcaggtt  | tttgttttca  | tcatccaaaa  | gtacttttta  | cagocaaagga | 4920 |
| cagtagcgaa  | ctttacatog  | catatgcaat  | gtgaagttct  | tgatcctcat  | cttacttaac  | 4980 |
| ctgtcagcag  | tatctgacac  | aggtgtcaot  | ggctcctccc  | tgagatgctc  | tctttatttg  | 5040 |
| gctttgggga  | caoccatatto | tcoccatctc  | tactttcctc  | aatggccctc  | ctcagctctc  | 5100 |
| tttggaasaga | ggcaaaaaga  | acttcattat  | ctcctggatg  | tagtacaaac  | aaotcaagct  | 5160 |
| caacatgtgc  | atactgaact  | tcoccaaact  | ttcccaaact  | tcgacattta  | cagccatccc  | 5220 |
| ctttcagctg  | atagcaagtt  | tatccttcca  | gctaactaaa  | ccagaatctt  | tagagccoatc | 5280 |
| cttgacccct  | ttcctcctct  | cacactcaac  | atctatccat  | cagaaaattt  | tggtgggtct  | 5340 |
| actttcaaaa  | gtcagagcat  | gtctcattac  | gtctcattac  | ctccaatagc  | taccatacta  | 5400 |
| gtctgaacaa  | acatcatttc  | tcactctggg  | tattgmaaaa  | acatcatttc  | tcacctgggt  | 5460 |
| tattgatagc  | atcctaaogg  | gtcttctctg  | ttcttgggtc  | ccctatatta  | gcacacacgc  | 5520 |
| agtcagagga  | gtccttttag  | aaactcaatc  | gacatgtca   | ogtcaotcct  | ctacttaaaa  | 5580 |
| tccttcaatg  | ggctcccatg  | acaaaacaga  | acaaaacaga  | gcctttacac  | tggtotacaa  | 5640 |
| gttccaacat  | ttgactcctg  | ttatctctct  | gacatcatat  | tctaataata  | ctgctgttgt  | 5700 |
| ccttttgtct  | agtcacacot  | gtttgattag  | taaatattta  | ttaaacaaag  | caatcctagt  | 5760 |
| ctccaaagag  | atcatagttt  | attggaggaa  | acaagagcct  | ataaatgggt  | acacacagaa  | 5820 |
| ggtagtgatt  | atggtttctc  | ctcactccc   | atcctaaaot  | ttgaacgggtg | aaactccct   | 5880 |
| ggatgttgaa  | gggttgaggaa | tttgccaggg  | ttcagggtgg  | tggttgaggga | ggcaggggag  | 5940 |
| aaacaaaggac | atttcaggca  | ggaagaaat   | tacatgcaaa  | gatctaaaga  | tatgaatcag  | 6000 |
| caacatattt  | atggaattac  | aagtaaaagta | gaaagtctct  | gctaanaaat  | caaaaaataa  | 6060 |
| agatttgtga  | ttagggggcc  | agaatgtggg  | agggaaagag  | agatacagtt  | cacactttta  | 6120 |
| gacaggagcc  | agatcatgaa  | atgttttctc  | tttgtttgtt  | tcttctctca  | cagottttga  | 6180 |
| tatgctcttg  | gagcaattta  | ttaaoccat   | tttttaatgc  | atctcctgaa  | cagagtcaaa  | 6240 |
| gcaatacttg  | gaaaggactc  | tgaatttctc  | gatttaaaag  | tacaaaagaa  | aaatctggag  | 6300 |



|             |            |            |            |            |            |      |
|-------------|------------|------------|------------|------------|------------|------|
| tcacaattaa  | tttgagaagg | taaaggagtg | gggtgtgtac | tgtatcaaat | ttaatttqta | 6360 |
| caaaatcatc  | atctctagta | acattatctt | ttctaattta | ctgcgtttag | actactttag | 6420 |
| taaagottga  | tctccctgtc | tatctaaaca | ctgattcaat | tacagcaagc | ttoaggotag | 6480 |
| cattgggtcat | attaataccc | aacaaatcca | caagggtgta | gttgcaaatg | atcttgtata | 6540 |
| aaagggtgaac | tgagatttca | ttcagtttac | agctcttgcc | aggcaaggca | gcagaccaca | 6600 |
| ggtaggtott  | ggcatctacc | gtttttaaag | gtgacagota | cttttgaaat | tacagatttg | 6660 |
| tcaggacatg  | gaggacaaaa | ctagagottc | tcactactgt | tgtgtaggaa | atctatgttt | 6720 |
| gtcaacctgg  | cttgtaaaat | atggttaata | taacgttaac | aotgttagca | agtaactgac | 6780 |
| tttatagacc  | aatatgcctc | tcttttgaaa | tggtottatt | ttaaacaaat | gtgagcaaaa | 6840 |
| gaaaatattt  | atgagattct | aaaaatgaag | acataatttt | gtagtataga | atcttcttgg | 6900 |
| ccagggaatg  | tggctcatgg | ttgtaatccc | agcaotttgg | gaggccaagg | tcagaggatt | 6960 |
| gcttgagcct  | gggaagggtg | agatgcagtg | atctaatggt | ataccactgc | actccagcct | 7020 |
| gggcaacaga  | gcaagaccc  | gtotoagaa  | aagaaaagaa | ttttatcttt | cttttcagac | 7080 |
| aaaaatagac  | tttaaaataa | taatggaaga | acaaatatga | tgtacacaa  | tatcagagta | 7140 |
| attactttat  | gacagttaag | ataagatttc | taattcttaa | atattcctct | gottaataca | 7200 |
| tttatattgga | gttttgatct | ataatatatt | cccaacctga | cccaaaaaat | gaagaaggac | 7260 |
| aaaggaaaaat | gttggtccaa | gaaacaaaga | tgtaagtaaa | aaggcaataa | gaaggaaaaa | 7320 |
| aaacttttga  | agcaaatgtg | gattgaggag | gatgagcaga | ccaatatttt | ttggtttggt | 7380 |
| cagottacat  | aatgattatc | ttcttttggg | ttctcagttt | ctagtgggtc | tcattgtttg | 7440 |
| cttccagac   | caggatgaag | acactccagt | ttttcttctc | ttctgtttgc | tggaagcaaa | 7500 |
| tctgctgcaa  | tagctgtgag | ctgacaaaca | tcacaaattg | aatagagaaa | gaagaatgtc | 7560 |
| gtttctgcat  | aagcatcaac | accaotttgg | gtgtgtgcta | ctgtacaccc | agggtaggta | 7620 |
| cc          |            |            |            |            |            | 7622 |

&lt;210&gt; 2

&lt;211&gt; 6038

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 2

|             |            |            |             |            |             |      |
|-------------|------------|------------|-------------|------------|-------------|------|
| ggatccgaga  | acatagaagg | agoaggtaat | ttatcaaggc  | atgaacacgg | gtgcttaatt  | 60   |
| tctatatttg  | aggccaggca | tggtgtgtca | caactgtaat  | cccaacactt | taggaagcca  | 120  |
| agtggtgttg  | attgottgag | tctaggattt | tgagaccagc  | ctggccaaca | tgggaaatc   | 180  |
| ctgtctctac  | taaaaataat | aaaattaaac | agtcattgtg  | gtggtgtgco | tttagtccca  | 240  |
| gtcactctgg  | tgggtgaggc | acaagaatca | cttgaaacctg | ggaggcagag | gttgcaagtga | 300  |
| gtgagactg   | tgccaactca | ctccagcctg | ggtgacagag  | taagattctg | tctcaaaaaa  | 360  |
| tattgtatata | tacacacata | taatagatac | ataaacatat  | atacatatat | aatatataaa  | 420  |
| tatatataatt | atatataata | tataaacata | tataaatata  | tatatatata | tatatatata  | 480  |
| tatatataacc | aaacataaag | gaataatttt | gggggaaaaa  | cttcataaat | gaagaaacaa  | 540  |
| cataggctgt  | tgagtatatg | cacagaaatt | caagagatct  | tccagcaatt | gaagacattg  | 600  |
| gtttactcaga | attcacaaaa | aagtcaagct | gtgcatttaa  | agtagaatgt | gatgagtgtt  | 660  |
| accactgagg  | taggaactgg | gaactaagga | agcgtaaagc  | agaaagtgtc | gaactgagag  | 720  |
| ttgggcattg  | gaggctgtgt | aaggcagggt | aagtgaatgt  | ctcctagaag | ctaactttta  | 780  |
| atggagtttt  | gaagtaottg | taggagtagc | ttaggtgaaa  | agangaggag | aaacatgtat  | 840  |
| caggcagagg  | gactagaacc | ttattacctt | caaaagaaga  | gcanaaagaa | tacatgtgac  | 900  |
| tttgagggtg  | tgggaggtgc | tttaagccaa | tataggtgaa  | tttgacatag | gaacttccct  | 960  |
| aataatgttc  | ggtcattttg | taaatattga | gtgatataat  | actgtattaa | agcccaagag  | 1020 |
| ttgtttttat  | atagaaagaa | gaaaaaagcc | caagagagtt  | ttatttctag | agggaatatt  | 1080 |
| ttotagaaat  | aaagggaagt | gtatcagcca | gtttctagtc  | aggaaaaacg | aaatcacacc  | 1140 |
| tgatatgcaa  | aatagaggaa | aatcagggaa | ttcattaatc  | oagagatttg | gttgctcaag  | 1200 |
| tattagattg  | ctgaaaagcc | agacagggsa | tatgaggcaa  | tcagagataa | gtattagtga  | 1260 |
| caagctccat  | ttatgtgcag | gattggaggg | acataggtgg  | ggttcccaga | agccagaagg  | 1320 |
| tgagaccacc  | tagcagaagg | tcacacacac | gotggggttt  | ctcacaacaa | gotgggacca  | 1380 |
| ccaggaggag  | ctgtccaatg | ggatctggag | ccaggagagt  | oatgcagtoa | ctaccaggaa  | 1440 |
| gggaagcaga  | atgtaaaagg | tagagagaaa | tactccaact  | gcttccttgc | attcactttc  | 1500 |
| caatctccat  | tcacaaaggc | aaaaacotgc | taatacagca  | gagtgggaaa | agcagcctgc  | 1560 |
| caaggctcct  | tctccacaaa | aacagagcac | aaaaaccaagc | aaaaacaaag | aatgcatttg  | 1620 |
| atagcaacaa  | ggctatggac | caacccaaca | taaaagaat   | gatgagtgat | ttcttttttc  | 1680 |
| atttggttoa  | agaaaagtat | ttcagtaact | attatgtaac  | agaaattota | tttatttttg  | 1740 |
| ggaattocaaa | ggtgaataaa | aaagaactct | aaatttttat  | caataaaata | tttcaaaaac  | 1800 |
| ctaatgaga   | gtaatggcat | taactagcaa | atatgtcaat  | gagatgagct | agocataaga  | 1860 |
| ggottagaat  | tgagagaaa  | gtctgggggc | ctcttgacag  | gcacaaatto | gagctgtttg  | 1920 |



|             |             |             |            |             |             |      |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| tggaatctc   | tgacctaaact | goaggtggaa  | atataaatat | gggcatttag  | aatagtggcc  | 1980 |
| caaaactttg  | atgattttctg | tcttggggtc  | totccaatta | atgggattga  | tgagaactgt  | 2040 |
| agaccaactg  | ggtcaccatg  | gotcaatgaa  | tagtcccctg | gctttggagt  | caaaactgacc | 2100 |
| tgaatatgaa  | ccocagcttt  | gotcaottaca | ggttgoattt | atcctcagtt  | ttctcatott  | 2160 |
| toaaagaaga  | acagtaactt  | otttaaaagg  | ttattgtagg | ctgggtgcag  | tggtcacgc   | 2220 |
| ctgtaatcgc  | agcaactttg  | gagggggagg  | ctagtggatc | aottgaggcc  | aggagttgga  | 2280 |
| aaactagcctg | ggcaacatgg  | tgaaactctg  | tctotacaaa | aagaaattta  | aaaaattttg  | 2340 |
| ctgggtgttg  | tggtcacacac | otggaaattc  | agctaccctg | gaggccgagg  | catgagcatc  | 2400 |
| acttgagctc  | ggaaagcaga  | gggttgagct  | gagocaaagt | tgtaccactg  | taotcaagcc  | 2460 |
| tggtgtgacac | agtgtgacct  | tgtctaaaaa  | aaaaaaaggt | tattgtgtta  | ttgtaaatat  | 2520 |
| tgtatatgaa  | cttctattta  | acatgtttag  | ttaaatgcct | gtgtaattgt  | ccaatgtgct  | 2580 |
| cttctagctc  | actgcacaga  | caaaaotgat  | tcactgaaat | catggaaattg | cagcaaaaga  | 2640 |
| caaatotaat  | taattgtaggt | caaacgggag  | gaotggagtt | attattotaaa | toagtctccc  | 2700 |
| tgaaaactca  | gaggctaggg  | ttttatggat  | aatttgggtg | gcaggggact  | aggggaatggg | 2760 |
| tgctgctgat  | tggttgggga  | atgaaatagt  | aagattgttg | aaaactgtcc  | tcottcattg  | 2820 |
| agtctgcttc  | cgggtgtagg  | ccacacgacc  | agttgagtc  | tgaagcatgc  | gtccaagtgg  | 2880 |
| agtcagtttg  | ttgcagaaat  | gcamaagcct  | gaaaaatgto | tcaaatgato  | aaotgtaggc  | 2940 |
| tccacaataa  | tgaattattc  | tataggagca  | attggggaag | taacaaatct  | tgtgacctct  | 3000 |
| ggacacataa  | ctcctgaact  | agtaagggat  | atgaaaaaac | atgootatat  | ottatoagaa  | 3060 |
| ttcaggtccc  | cccataatcc  | taatctocac  | goatttcatt | tgtttagaaa  | ggcoattttc  | 3120 |
| agtccctgag  | caaggagggg  | gttagtttta  | ggataggact | attatoottg  | cttctgtaaa  | 3180 |
| ctataaacta  | atttctctcc  | atggttagct  | tggtctacac | ctaagaatga  | gtgagaacag  | 3240 |
| ccagcctgtg  | aggctagagg  | caagatggag  | tcagccatgc | tagattttatc | toactgtcat  | 3300 |
| aaoccttgoa  | aaggcagttt  | cacctgggac  | ataggaggta | ctcaatgaaa  | aagaagctat  | 3360 |
| taatatataa  | attttaaaaa  | tgaatttaag  | gaactaatac | tatgtacata  | ttagtcatat  | 3420 |
| aaacaagtg   | gttcattttac | atcaacacaa  | ataaatcttg | tgattatata  | taggtaatat  | 3480 |
| gaaaaacttt  | gtttctcttc  | ataatacaag  | gtattagcaa | tagatatagt  | aatgttagoa  | 3540 |
| ttccttttga  | aaaaatgaaa  | agattttata  | tttcccaaga | atoatttagta | tttttattta  | 3600 |
| atatacatata | tataaaattt  | attcatctta  | taaatgtgaa | atatgcttgc  | ttacaaatta  | 3660 |
| ctgacagatt  | tczaaatatt  | totatactoa  | caattatcat | ttacataaat  | attgatttgg  | 3720 |
| taottacaat  | gtgtactgct  | atgctaagtt  | ttgtctttgt | caaacatatt  | ttataaaatc  | 3780 |
| ataatcctag  | atgaatooag  | cttttggtaa  | cccagctgoc | tgaacccctg  | ctgttaacag  | 3840 |
| gcaagtggtg  | gtagggtacag | atctatacct  | accaoottoc | tctacccacc  | agcatctgca  | 3900 |
| cccaacaccc  | ctcccaaccc  | accattatct  | ataccaacca | ccoctcccaa  | cctaocagca  | 3960 |
| tctgcaccca  | ccacacogcc  | cacccaccc   | catgtacact | cactacacct  | tcagccatc   | 4020 |
| acatctgca   | cccatcactc  | ctcccactcc  | acaagcatct | gcauccaoca  | catttcccta  | 4080 |
| cctacacagca | tcttcaactc  | ccacotctcc  | accacacagc | atctgcaccc  | accaacccctc | 4140 |
| ctcaacacac  | agagtctgca  | tocatcacac  | ttgcccactc | gctagcatct  | gcacacatca  | 4200 |
| gctctgcctt  | cttgcctaat  | acgggatgag  | ctctccatgg | ttctgcctaa  | agacaaatgct | 4260 |
| tcacactctc  | tctctataacc | acttctctt   | tactcttca  | agtaacatc   | agaacttctc  | 4320 |
| tctctctctg  | ataccaaactt | tttccactt   | actcaatcat | tctatcaccc  | atacaaaagt  | 4380 |
| gtttatctct  | cccatcttaa  | agttaaaaat  | caaaagaaaa | ttgtctgggy  | ccaggcaagg  | 4440 |
| tggtctacgc  | ctgtaatccc  | aaactcttgg  | gaggccaagg | aggggttggt  | gacttaaggt  | 4500 |
| taggagttoa  | agacacagct  | ggocaaacatg | gtgaaaccca | tctotactaa  | aaatacaaaa  | 4560 |
| attagccagg  | catggtggca  | catgcctgta  | gtctcaggta | cttgggaggc  | tgaggccaga  | 4620 |
| gaatggcttg  | aaocccgggag | gcagaggttg  | cagtgtgagc | agattgtgoc  | cttgcactcc  | 4680 |
| agcctgggtg  | acagagtgag  | actocatctc  | aaaaataaaa | aatataaata  | aaacaaaaga  | 4740 |
| aagttaattt  | tacccaacat  | ccacattaac  | caaataccca | ttcttttatt  | gatctttgta  | 4800 |
| aaaaaaagct  | cttgggaaaa  | ttgtctatat  | tcactatgac | ttatctcttc  | caaatcaott  | 4860 |
| aaacacatac  | caatcaggtt  | tttgttttca  | tcattccaaa | gtacttttta  | cagccaagga  | 4920 |
| cagtagcgaa  | ctttacatcg  | catatgoatt  | gtgaagttct | tgatctctat  | ottacttaac  | 4980 |
| ctgtcagcag  | tatctgacac  | aggtgtcact  | ggctcctccc | tgagatgtct  | tctttatttg  | 5040 |
| gctttgggga  | cacocatattc | tcccacttcc  | tactttctct | aatggccctc  | ctcagttctc  | 5100 |
| tttggaagaa  | ggaaaaagaa  | aottcattat  | ctcctggatg | tagtacaac   | aactcaagct  | 5160 |
| caacatgtgc  | atactgaact  | coatttctct  | ttcccaact  | tcgacattta  | cagccatccc  | 5220 |
| ctttcagctg  | atagcaagtt  | tatccttoca  | gctaactaaa | ccagaatott  | tagagccatc  | 5280 |
| ottgacoott  | ttctctctct  | cacactcaac  | atctatccat | cgaaaaattt  | tgttgggtct  | 5340 |
| actttcaaaa  | tgcatcacaga | gtcagagcat  | gtctcattac | ctcaaatagc  | taacatacta  | 5400 |
| gtctgaacaa  | acatcatttc  | tcactgggtt  | tattgaacaa | acatcatttc  | tcactgggtt  | 5460 |
| tattgatagc  | atcctaacgg  | gtcttctgt   | ttcttgggtc | ccotatatta  | gcacacagc   | 5520 |
| agtcagagga  | gtccttttag  | aactcaatca  | gatoatgtca | cgtoactcct  | ctacttaaaa  | 5580 |
| tcttcaatg   | ggtcccatat  | cacaaagagt  | acaaaacaga | gcccctacac  | tggtctacaa  | 5640 |



|             |            |             |            |            |            |      |
|-------------|------------|-------------|------------|------------|------------|------|
| gttccaaacat | ttgactootg | ttatototot  | gacatoatat | totaatatta | otgotgttgt | 5700 |
| cotttttgto  | cagtcacact | gttttgattag | taaatattta | ttaaacaaag | caatcctagt | 5760 |
| ctccaaagag  | atcatagttt | attggaggaa  | acaagagoot | ataaatgggt | acacacagaa | 5820 |
| ggtagtgatt  | atggtttccc | ctcacotccc  | atcctaaact | ttgacaggtg | aaactccoct | 5880 |
| ggatgttgaa  | ggttgaggaa | tttgccaggg  | ttcaggggtg | tggtggagga | ggcagggagg | 5940 |
| aagcaaggac  | atttcaggga | ggaagaacat  | tacatgcasa | gatotaaga  | tatgaatcag | 6000 |
| caacatattt  | atggaattac | aagtaaaagta | gaaagttc   |            |            | 6038 |

<210> 3  
 <211> 542  
 <212> DNA  
 <213> Homo sapiens

|            |            |             |            |            |             |     |
|------------|------------|-------------|------------|------------|-------------|-----|
| <400> 3    |            |             |            |            |             |     |
| tcaotgttag | caagtaactg | actttataga  | ccaatatgcc | totottotga | aatgggtotta | 60  |
| ttttaaacaa | atgtgagcaa | aagaaaatat  | ttatgagatt | ctaaaaatga | agacataatt  | 120 |
| ttgtagtata | gaattttctt | ggccaggaat  | ggtggctcat | gottgtaate | ccagcaotttt | 180 |
| gggaggccaa | ggtcagagga | ttgcttgagc  | ctggaagggt | gaagatgcag | tgattcoatga | 240 |
| ttataccact | gcactccago | ctgggcaaca  | gagcaagacc | ctgtotcaag | aaaagaaaag  | 300 |
| aattttattt | ttcttttcag | acaaaaatag  | actttaaaat | aaataatgaa | gaacaaatat  | 360 |
| gatgatcaaa | attatcagag | taattacttt  | atgacagtcg | gaataagat  | totaatottt  | 420 |
| aaatattoct | ctgottaaat | catttatattg | gagttttgat | ctataatata | ttcccacott  | 480 |
| gaoccaaaaa | ttgaagaagg | acaaggaaaa  | atgtttgttc | aagaaacaaa | gatgtaaagta | 540 |
| aa         |            |             |            |            |             | 542 |

<210> 4  
 <211> 3213  
 <212> DNA  
 <213> Homo sapiens

|             |             |             |             |             |            |      |
|-------------|-------------|-------------|-------------|-------------|------------|------|
| <400> 4     |             |             |             |             |            |      |
| actaacataa  | agctgaagg   | gaataaaaaa  | atcagggtta  | gccaaacaaa  | ttttcatggt | 60   |
| caaataccac  | ataaaaagta  | aatatactta  | agttccagc   | aaaatctgaa  | ttgaacgtag | 120  |
| acaaaatgot  | cattttctcag | tgtttgacag  | acttaacagt  | ttgagccaat  | aaaaatgtac | 180  |
| tgactagata  | aactactaaa  | agttgttaat  | ttttgcaatg  | tatatctctg  | aaaagaaagt | 240  |
| ttatotatta  | tagaaattcc  | tggtcccat   | taagaaacttt | gagcattttta | attgtttaat | 300  |
| aatatagttt  | aattgcacaa  | tgaaaataat  | caataataca  | atttatttgg  | tttattttaa | 360  |
| aaaactgatt  | ottttctgctc | ctctatatata | tagaotgatt  | ttataactat  | gttgootaaa | 420  |
| gatoaccaaa  | ttgtttgaag  | cctagggttc  | tgagggtagg  | aaaatgatgt  | cccaactatt | 480  |
| tacagttcac  | acacacatto  | tggtgtttta  | atacatcctt  | tacaaagtga  | ggaaaggtgg | 540  |
| aagattgatg  | atttggggga  | attagagcta  | ccacacccca  | gagggtggta  | tggtatgttg | 600  |
| totgttgtga  | gctgtgtgaa  | tcagagagtt  | tgatttagac  | atatatttag  | aaagaggaaa | 660  |
| gatgaaccaa  | tcaaaaataa  | taactataat  | gacttttcaa  | gatatagaac  | atacagttaa | 720  |
| gatataaatg  | gaaacaaaaa  | aagttaaaag  | tggtggagatg | aagtotgatt  | ttttggtttt | 780  |
| tttttttttt  | tgcttttttg  | tttgtttatg  | taatcagtg   | tacpagttta  | aaataatggg | 840  |
| ttataagaac  | ctatatgcac  | gootcatggt  | aaotccaat   | ctaaaacata  | caaacataac | 900  |
| acacaaaata  | aaaaggagaa  | attaaaacac  | accaccagag  | aaaatcacct  | acattaaaag | 960  |
| aaagacaaat  | aggaagaaaa  | taagaagaag  | aaggccatca  | aatatcaga   | aatgaataa  | 1020 |
| caaatgaca   | ggaaatagtc  | ctcataaata  | ataacattga  | atgtaaatgg  | actaagotct | 1080 |
| ccaatgaaag  | acagggagtg  | gctgaatgta  | ttttaaaaaa  | aatattacac  | cgagctgtgc | 1140 |
| gtggtgtoto  | acacotataa  | tccagcagatt | ttgggagact  | gagccgggtg  | gatoacttga | 1200 |
| gcccaggagt  | tcgagaccag  | cctgggcaac  | atggcaaaac  | cotgtctcta  | ctaaaaatag | 1260 |
| aaaaaattag  | ctgaacatgg  | tggaacatgc  | ctgtggttcc  | agctactaga  | gaggctgagg | 1320 |
| oagaagaatt  | gottgaactt  | gggaggtgga  | gggtgcagtg  | agotaagatt  | gatggagoc  | 1380 |
| ctgcacccca  | gcctaggtga  | cagaataaga  | ctctgcctca  | aaaaaaaaaa  | gcaaaacaaa | 1440 |
| acaaaacaaa  | aaacccttag  | acccaatgat  | tcattgucata | caagaagtat  | gcttccactt | 1500 |
| taagacacaa  | tatagactga  | aggtaaaggg  | atggaaaaat  | attctatgcc  | tatggaacaa | 1560 |
| acaaaaaaga  | agcagaagot  | acattttatat | cagacaaaat  | agactgcaag  | acaaaaacta | 1620 |
| tgaaaagaga  | gaaagaaggt  | catttatatag | tgataaaggg  | gtocatttag  | caagagcatt | 1680 |
| taacaattct  | aaatatatat  | tcacccaata  | ctggagtaact | caggtatata  | aagcaaatat | 1740 |
| tattagagcc  | aaagagagag  | atagaagagc  | cccatacaa   | taataactgg  | agacttcaac | 1800 |
| aoocccacttt | cagcattgga  | cagatcatcc  | agaacagaaa  | ttacaaacaa  | tcaaatttca | 1860 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| totgoacocat | aggtomaatg  | gacotagtag  | atatttaoag  | aacatttgat  | ccaacagotg  | 1920 |
| tagaataoac  | attcttctco  | tcagcacatg  | gataattctc  | aaggatatac  | caaatgtag   | 1980 |
| gtcacaaaac  | aaatcttaaa  | atthagaaaa  | aaagtgaat   | aatatcaaac  | gttttctotc  | 2040 |
| accacagact  | aagaaaaaaa  | gaagtcccaa  | ataaatacaa  | tctgagataa  | aaaaggagac  | 2100 |
| gagacaacca  | ataccaaaaa  | aaattaagg   | atcattagaa  | gatactatga  | aactatgatg  | 2160 |
| taataaattg  | gnaaacctga  | acaaaataga  | taattcootag | aaacatacaa  | catactggto  | 2220 |
| tgtttaggtt  | ttgtattttt  | tcattagtacc | atgaagaaat  | acaagaattg  | tttctagaac  | 2280 |
| oattotttga  | tttcttoatg  | gtttttgtat  | ttcttoatgg  | aaootgaag   | aaatacaaaa  | 2340 |
| tgtgaacagg  | ccaataacaa  | gtaatgagao  | agaagccata  | ctaaaaagta  | tcccagaaaa  | 2400 |
| gaactcagga  | tctgatggot  | tcactgatga  | attttgocaa  | atatttaaaa  | aaotaatacc  | 2460 |
| aatocaaotc  | aaattatttaa | aaaaatagag  | gtggacagaa  | tottttccaaa | tgtattctat  | 2520 |
| gaggccagtg  | ttttttctga  | ttgaatctcc  | cattatattt  | taatoacata  | taaaaocaga  | 2580 |
| gaaagacaca  | ttaaaaagaa  | agaaaaotgt  | aggccaatat  | ctotgatgaa  | cattgatgca  | 2640 |
| gaaatcctca  | acaacaaatt  | agcaaatga   | attcaagaac  | acatttaaac  | aatcattcoat | 2700 |
| catgacocagg | tgttaatttg  | oatagagatt  | caagtgtggt  | taggtatgtg  | cagatoaatg  | 2760 |
| gggttaaatgt | tgtccaatga  | acataatgtc  | ctccagotcc  | atccatgttc  | ttgcaaatga  | 2820 |
| caggatctca  | ttctttttta  | tggctaagta  | gtactcoatt  | gtgtataagt  | gcataatttt  | 2880 |
| otttatccat  | tcactctgta  | gacaocaaag  | ttgottccaa  | atotttagota | ttgtgaatag  | 2940 |
| tgtctgaata  | aacatgggag  | tgtaaatatt  | ttgttgacat  | actgatttca  | tttcttttgg  | 3000 |
| ataaataccc  | agttagtgga  | ttgctggatc  | atatggggga  | aatggagat   | ggctaacggg  | 3060 |
| ctcaaaaata  | tagttagaaa  | aaatgaatat  | gatttagtat  | tcgatagcac  | aataggatga  | 3120 |
| ctactgttaa  | tgataattta  | taaaataact  | aaaatagtat  | aaaatagtat  | aatgggatg   | 3180 |
| tatgtagcag  | agagaaatga  | taaattgttg  | aag         |             |             | 3213 |

&lt;210&gt; 5

&lt;211&gt; 6679

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 5

|             |             |             |            |              |             |      |
|-------------|-------------|-------------|------------|--------------|-------------|------|
| gtcgacctgc  | aggtcaacgg  | atcaottgag  | gacagttagt | caagaccago   | ctggggcagca | 60   |
| tagggagact  | gtctctacga  | aaaatoaaaa  | aattatggco | gggcattgggtg | gtccacgtct  | 120  |
| gtaatccctg  | aacttttgga  | catcaaggca  | agtggatcac | ttgaggtoag   | gagttogaga  | 180  |
| ctagcctggc  | caacatgggt  | aaacotatc   | tcactaaaa  | aatcaaaaa    | ttagccaggc  | 240  |
| atggtggcag  | gcacctgtaa  | tcocggotac  | tcaggaggot | gaggcaggag   | aatcaottga  | 300  |
| accagggagg  | cggagggttg  | agtgaotga   | gatoacacca | ctgcactcca   | gcctgggtga  | 360  |
| cagagcaaga  | ctctattctc  | aaaaaaataa  | aaaaataaaa | aaattagcca   | ggcatggtag  | 420  |
| tgoacacctc  | tagtctcagc  | taotcaggag  | gctgagggtg | gaggatocac   | tgaacctggg  | 480  |
| gcagtcagg   | ctacagtgag  | ccaagatcat  | gcactacac  | tcagagctgg   | gcaacagaga  | 540  |
| gagacccctg  | ctctaaaaaa  | ataataataa  | taagaaaaaa | aacagctctg   | tttatgtctc  | 600  |
| ctggtccata  | catactacta  | tgtatatagt  | ttgcaaaact | aaagatocag   | atagtoaat   | 660  |
| ttttaggott  | gtggggccgt  | tgggtctctg  | cacaatcaot | otgccotgtc   | tttctagcac  | 720  |
| aaaagcagct  | ataaacaata  | catacatgaa  | ttttttatag | acatcagat    | ttgaatttca  | 780  |
| tatgattttt  | acatttttata | aaataatott  | tttaaaaaat | ttccoctaac   | cattttaaag  | 840  |
| tgtaaaagoo  | ggccagggcg  | ccatcgtcac  | gcotgtaatt | ccagcacttt   | gggaggctga  | 900  |
| ggtgggcaga  | tcacttgaga  | tcacacagttc | gagaccagcc | tggccaaat    | agcaaaaacc  | 960  |
| cattctctact | aaaaataaaa  | aaattagctg  | ggcatagtgg | tgcacacctg   | tgatocacgc  | 1020 |
| tacttgggag  | gctgaggcag  | gagaatogct  | tgaaacotgg | aagcggagggt  | tgcagtgaac  | 1080 |
| caacatcatg  | coacotgact  | coagcotggg  | tgacagagtg | agaottogtc   | tcaacgaaaa  | 1140 |
| aaaaaagtgt  | aaaagccatt  | cctaattcag  | tgtacatcag | tgtacatact   | cagggtotgog | 1200 |
| tactcctgct  | ctgaggcata  | cctgagagtg  | agagttgott | ggtcacaggga  | catcacacatt | 1260 |
| tcacatttaa  | ctagacacta  | ccaagtggcc  | atccaaggag | gttttttttt   | tacaatctac  | 1320 |
| actcccccca  | gcaacaaatg  | agagttactc  | cagatocctt | acaaagatgc   | tctaagccca  | 1380 |
| gtaccagatg  | aaaacaggaa  | gtgggagggg  | aagctgccag | occttcttaa   | ccatgaagaa  | 1440 |
| atacctggta  | gagccttctg  | ggatgtggaa  | ggatgaataa | cgggggtctc   | tggagootgc  | 1500 |
| ccootgtog   | atoactgtga  | ottctgagcc  | tcagctocag | tctcagcccc   | atgtgtcatg  | 1560 |
| gccagtata   | atgagccctc  | actctctgtt  | tggcttttat | tctccoccatg  | tggggctgaa  | 1620 |
| gtotggattg  | agccgttatt  | caagatgtac  | agctttcttg | acaggaaagt   | agtgtcacag  | 1680 |
| aaacagcagg  | ggcttggoaa  | gatgatctaa  | ctgoaaatcc | tacotggotc   | agccaccago  | 1740 |
| tagttotgtg  | atcttgaaac  | agttttttca  | cttctctgag | gccatccctt   | ggctacacaa  | 1800 |
| caacagttgg  | ttgacaggat  | gaaatgacga  | agtcocctac | acotgtaatc   | ccagocacttt | 1860 |
| gggaggcoaa  | ggcgggtgga  | tggcttgagc  | ctgagaggtg | acagcatgoc   | ggcagctctc  | 1920 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| acagccctcg  | ttgctctctg  | ggcctctctc  | tgctgggggt  | cccacttogg  | tggaacttga  | 1980 |
| ggagcccttc  | agccacacgc  | tgactctgtg  | gagccctttt  | ctgggctggc  | caaggccaga  | 2040 |
| gocggctccc  | tcagcttgca  | gggaggtgtg  | gagggagagg  | ctcaagcagg  | aaccgggggt  | 2100 |
| ggccacggcg  | cttgccgggc  | agctggagtt  | ccgggtgggc  | gtgggtttgg  | ogggcccccgc | 2160 |
| actcggagga  | gogggccagc  | cctgcacagg  | cccgggcaat  | gagaggttta  | gcaccggggc  | 2220 |
| cagcggctgc  | gaggggtgta  | ctgggtgccc  | cagcagtgoo  | agcccgccgg  | cgtgtgtctc  | 2280 |
| gctcgatttc  | tcactgggco  | ttagcagoot  | tcocgggggg  | oagggctcgg  | gacctgcagc  | 2340 |
| cgcacatgco  | tgagcctccc  | ctccatgggo  | tcctgtgcgg  | cccgagcoto  | cccgacgagc  | 2400 |
| acccacccct  | gctccacagc  | gocccagtccc | atcgacccag  | caagggctga  | gaagtggggg  | 2460 |
| cgcacgggac  | cgggactggc  | aggcagctac  | ccctgcagoc  | ctgggtggga  | atccactggg  | 2520 |
| tgagcgcagc  | tgggctcctg  | agctctgtgg  | agaottggag  | aacctttatg  | totagotcag  | 2580 |
| ggatcgtaaa  | taacccaato  | agacccctgt  | gtotagotca  | gggtctgtga  | atgcaccaat  | 2640 |
| ccacactctg  | tatctagota  | ctctgatggg  | goottgagga  | aocctttatg  | ctagotcagg  | 2700 |
| gattgtaaat  | acaaccaatc  | goactctgtg  | totagotcaa  | ggtttgtaaa  | caacccaato  | 2760 |
| agacccctgt  | gtctagctga  | gggtatgtga  | atgcaccaat  | cgacagctctg | tatctggcta  | 2820 |
| ctttcatggg  | catccgtgtg  | aagagaccac  | caaacaggct  | ttgtgtgagc  | aataaagctt  | 2880 |
| ctatccactg  | ggtgcagggtg | ggctgagtoo  | gaaagagag   | tcagcgaagg  | gagataaggg  | 2940 |
| ctggggcogtt | ttataggatt  | tgggtaggtg  | aaggaaaatt  | acagtcacaa  | gggggttgtt  | 3000 |
| ctctggcggg  | oagggagtgg  | gggtcgcaag  | gtgctcagtg  | gggggtgctt  | ttgagccagg  | 3060 |
| atgagccagg  | aaaaggactt  | toacaaaggt  | atgtoatcaa  | ttaaaggcaag | gacccggcat  | 3120 |
| ttacacccct  | tttgtgtgtg  | aatgttoatc  | gttaagttgg  | ggcagggcat  | attcaottct  | 3180 |
| tttgtgattc  | ttcagttact  | toagggcoato | tgggcgtata  | tgtgcaagtt  | acaggggatg  | 3240 |
| ogatggcttg  | gcttggtctc  | agaggtctga  | cagctactct  | ggtggggcct  | tgagagaatg  | 3300 |
| ttgtgtcgao  | actctgtatc  | tagttaatct  | agtggggaog  | tgagagaacct | ttgtgtctag  | 3360 |
| ctcagggatt  | gtaaacgcac  | caatcagcgc  | cctgtcaaaa  | cagacccactc | ggtctatcca  | 3420 |
| atcagcagga  | tgtgggtggg  | ccagagataag | agaaataaag  | cagggtgccc  | gagccagcag  | 3480 |
| tggaacgcgg  | oacaggtccc  | tatccacaaat | atggcagott  | tgttcttttg  | ctgtttgcga  | 3540 |
| taaatcttgc  | tactgctcgc  | tttttgggtc  | oacactgott  | ttatgagctg  | taacactcac  | 3600 |
| oacgaaggto  | tcagagotto  | ctcctgaago  | caactaagaoo | acgagccccc  | cgggagggaat | 3660 |
| gaacaaactc  | ggccgcgctg  | occtaaagag  | tataaactoc  | acccggaagg  | totgcagott  | 3720 |
| oactcctcag  | ccagccgagac | cacgaacocaa | ccagaaggaa  | gaaaactgoga | acacactctga | 3780 |
| acatccagaag | gaacaaactc  | cagatgcacc  | accttaagag  | ctgtaacact  | oactgogagg  | 3840 |
| gtcccgccgt  | tccttcttga  | agtoagtgag  | acacagtttc  | cacacacagg  | ggacacaaag  | 3900 |
| ccaggagttt  | gagatcagcc  | tgggcaccat  | gatgaaatgc  | octctctgoa  | aaaaaanaaa  | 3960 |
| aaattacaaa  | aatctggcga  | goatggtggt  | cogtgccctgt | ggtccacagct | acccgggagg  | 4020 |
| ctaaagtggg  | aggatogctt  | gagcctggga  | ggtgaagact  | gcagtgagot  | gtgattgtac  | 4080 |
| ccacgcccct  | taggctgggg  | gacagactga  | gaocctgttt  | ccocctcoga  | aaaaaattga  | 4140 |
| caaaagtgtg  | ataagagggtg | cctgatatgg  | ctaggccgag  | tggtctcatgc | ctgtaatccc  | 4200 |
| agcacttttg  | gaagccgagg  | ogggcgggtc  | acctaaggct  | aggagtgtga  | gaocagocctg | 4260 |
| gocaaacatg  | agaaagccca  | totctcttaa  | aaatacaaaa  | ttagccgggt  | gtgggggocag | 4320 |
| tggtggagga  | tgctgttaat  | ccagotactc  | ccaggaggtg  | aggcagggaga | atcacttgaa  | 4380 |
| ccacggaggc  | ggcgggttga  | gtgagccgag  | atogtgccat  | tgcactccac  | ccactccagc  | 4440 |
| ctgggcaaca  | agagccaaac  | tcctgtctaa  | aaaaaanaaa  | aaaaagtgc   | tgacatataa  | 4500 |
| gaggtgtgga  | atgcaaatagt | tgccaggcaa  | oatgtttaag  | aatgtggago  | tcotgoottc  | 4560 |
| catggtcctg  | ttaaaaaccc  | acccccaagg  | ccagggtgcag | tggtctcatgc | ctataatccc  | 4620 |
| agoacttttg  | gagggccgagg | oggggtggatc | aoctgaggtc  | aggagtctga  | gaccagccctg | 4680 |
| accaccaaoc  | tggtgaaatc  | ccacotctac  | taaaaaataca | aaattagatg  | agcatggtgg  | 4740 |
| tgoatgcctg  | taattcccac  | tacttgggag  | gctgaggcag  | gaaaatcaot  | agaaccaggg  | 4800 |
| aggcggagggt | tgtagtgagc  | ogagatcgtg  | coattgcact  | ccagcctgag  | caatgagoga  | 4860 |
| aactccatct  | caaaaaaaca  | acaacaaaaa  | ccactctctc  | actcccaggg  | agctgggtac  | 4920 |
| agagctgggg  | cacatcagtg  | caagggtgctg | agccacagag  | otaaggogga  | gotgagggac  | 4980 |
| cgcggaccag  | ataacagtg   | gtgagatcag  | tggtgtgagat | cagacgtccc  | tgocattggg  | 5040 |
| gaccaccagg  | ggggcccccac | goacccagaga | tgcccccato  | cagtcacccac | atccacttct  | 5100 |
| catccagaga  | tgtctgtttc  | ttggcaogot  | ggggtaaatt  | aggacagaa   | gtgacagtot  | 5160 |
| tggtgtgtgt  | cagtcagact  | gcccagggca  | ggccttgggg  | octgtagaaa  | acgttoagge  | 5220 |
| ctaggccggg  | ccgggtggot  | caogootgta  | atcccaagao  | tttgggaggc  | cgaggcgggt  | 5280 |
| ggatccagag  | gtcagggagat | cgtgacccatc | ctgggtacaa  | cgggtgaaac  | cogtctctao  | 5340 |
| taaaaaatac  | aaaaattggc  | ogggcaatggt | ggcggggcacc | tgtagttcca  | gotactcggg  | 5400 |
| aggotgaggo  | aggagaatgg  | ogtgaaacccg | agaggccagag | tttgaggtga  | gocgagatcg  | 5460 |
| ogccactgca  | ctccagcctg  | ggogacagag  | caagactoca  | tctggaaaa   | aaaaagaaaa  | 5520 |
| ogttcagggtc | tgagccagag  | gcccagggtg  | taattctgtc  | acttaacatg  | accttgggga  | 5580 |
| aggcaottcc  | ttccctgggc  | oagttcacgg  | ggttggaato  | gaotccaaag  | tcctctccag  | 5640 |



|             |            |            |            |            |            |      |
|-------------|------------|------------|------------|------------|------------|------|
| catttaacgct | gcatgggtct | aagatgagaa | gatggggcag | tttccctct  | ctcaccacag | 5700 |
| ccogtgtcca  | cttcaagggt | aatgaccagg | gaagtacagt | gtcccaatcc | cgcagttcca | 5760 |
| aagcccttgg  | ggacccctac | tgacgggtcg | tgacagagg  | ggtgaaggto | aggtgagcca | 5820 |
| atcgccctga  | aggggtcttg | ctcattcggt | acagacatcc | gggttctct  | gggtctaccc | 5880 |
| ggattctagg  | ggcttttagc | gaatgagtc  | tggggggcgg | gggggtttct | gggggagttc | 5940 |
| ccagctaatc  | aaacttggga | aggacagcct | ggaactttcg | atggtgctta | tccaaagtgt | 6000 |
| gggtgggac   | agcagccaag | acccaatgtc | ottatctcag | gtaggggctc | aggaggtctc | 6060 |
| ccagacaggg  | agcctccgga | gagtttgggg | gtagggaatg | gagcaaccag | gottcttttt | 6120 |
| ttctctctta  | gaatttgggg | gotttggggg | caggcttgag | aatcccaaag | gagaggggca | 6180 |
| aaggacaact  | cccacaaagt | ctgcccagag | gagagagggg | gaccccgaat | cagctgocac | 6240 |
| ttccccacag  | gcctctgcgg | cttccaggcg | ttatctcagc | gctcagcctt | tgttcagctg | 6300 |
| ttctgttcaa  | acaactctgg | gccattccag | cctgggtggg | gcagcgggag | gaagggagtt | 6360 |
| tgaggggggc  | aaggcgagct | caaaaggagg | tcagagattc | cacaatttca | caaaaacttc | 6420 |
| gcaaaacagt  | ttttgttcca | accccccctg | attgtcttgg | acaccaaatt | tgcataaatc | 6480 |
| ctggggaagt  | attaactaag | cttagtctgt | gccccaggta | atttccctcc | aggcctocat | 6540 |
| ggggttatgt  | ataaaggggc | ccctagagct | ggggccccaa | acagcccgga | gcctgcagcc | 6600 |
| cagccccacc  | cagacccatg | gctggacctg | ccaccacag  | cccatgaag  | ctgatgggtg | 6660 |
| agtgtcttgg  | cccaggatg  |            |            |            |            | 6679 |

&lt;210&gt; 6

&lt;211&gt; 6235

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 6

|            |            |            |            |             |            |      |
|------------|------------|------------|------------|-------------|------------|------|
| gatcacttga | ggacagtagt | tcaagaccag | cctggggcag | ataggagagc  | tgtctctacg | 60   |
| aaaaatcaaa | aaattatggc | cgggcatggt | ggctoacgtc | tgtaatccct  | gaactttggg | 120  |
| acatcaaggc | aagctggatc | cttgaggtoa | ggagttcgag | actagcctgg  | ccaacatggt | 180  |
| gaacccctat | ctccactaaa | aaatacaaaa | attagccagg | catgggtggc  | ggcacctgta | 240  |
| atcccggtta | ctcaggaggg | tgaggcagga | gaatcacttg | aaccacaggag | ggggaggttg | 300  |
| cagtgagctg | agatcacacc | actgcaactc | agcctgggtg | acagagcaag  | actctatctc | 360  |
| aaaaaaata  | aaaaaataaa | aaaatttagc | aggcatggta | gtgcaacact  | ctagtctcag | 420  |
| ctaactcagg | ggctgaggtg | ggaggatcac | ttgaacctgg | ggcagtcagg  | gctaagtgta | 480  |
| gccaagatca | tgccactaca | ctccagcctg | ggcaacagag | agagaccctg  | tctctaaaaa | 540  |
| aataataata | ataaagaaaa | aaacagctct | gtttatgtct | cctggtccat  | acatactaot | 600  |
| atgtatatag | tttgoaaaa  | gaagatcca  | gatagtcact | tttttagggt  | tgtgggccgt | 660  |
| atggtctctg | tcacaatcac | tctgcccctg | otttctagca | caaaagcagc  | tataaacaat | 720  |
| acatacatga | attttttata | gaacatcgag | tttgaaattc | atatgatttt  | tacattttat | 780  |
| aaaataatct | tttttaaaa  | tttcccttaa | ccattttaa  | gtgtaaaagc  | cggccagcgc | 840  |
| gccatogtca | cgccctgta  | tcagcactt  | tgggaggctg | agggtggcag  | atcacttgag | 900  |
| atcaacaggt | cgagaccagc | ctggccaaca | tagcaaaacc | ccatttctac  | taaaaataaa | 960  |
| aaaattagct | gggcatagtg | gtgcacacct | gtgatccag  | ctacttggga  | ggotgaggca | 1020 |
| ggagaatcgc | ttgaacctgg | gaagcggagg | ttgcagtgag | ccaacatcat  | gccaotgcac | 1080 |
| tcacagcctg | gtgacagagt | gagacttcgt | ctcaacgaaa | aaaaaaagt   | taaaagccat | 1140 |
| tcctaatcca | gtgtacatca | gtgtacatac | tcaggtctgc | gtaactcctg  | tctgaggcat | 1200 |
| acctgagaag | tagagtgtgt | tggtcacagg | acatacacat | ttccaactta  | actagacact | 1260 |
| accaagttgc | catccaagg  | ggtttttttt | ttacaactca | cactcccccc  | agcaacaaat | 1320 |
| gagagttact | ccagatcctt | taaaaagatg | ctctaagccc | agtaccagat  | gaaaacagga | 1380 |
| agtgggaggg | gaagctgcca | gccctttcta | accatgaaga | aatacctggt  | agagccttct | 1440 |
| ggatgctgga | aggatgaata | acgggggtct | ctggagcctg | ccccctgtoa  | gatcaactgt | 1500 |
| aottctgagc | ctccagtcga | gtctcagccc | catgtgtcat | ggccagtgat  | aatgagccct | 1560 |
| caotctctgt | ttggtottta | ttctcccoct | gtgggggtga | agtctggatt  | gagccgttat | 1620 |
| tcagagatga | cagctttctt | gaaggaaaag | tagtgtcaca | gaacacagag  | gggcttggca | 1680 |
| agatgatcta | actgcanaac | ctacctgggt | cagccaccag | ctagtctctg  | gatcttgaac | 1740 |
| aaagtttttt | acttctctga | ggccatccct | tggctaacac | acacacagtg  | gttgacagga | 1800 |
| tgaaatgacg | aagtccctta | caactgtaat | cccagcactt | tgggaggcca  | aggcgggtgg | 1860 |
| atggcttgag | cttgagaggt | gacagcatgo | cgccagtcct | ccagacccct  | gttcgctctc | 1920 |
| ggcgccctct | ctgcccgggc | tcacactctg | gtggcacttg | aggagccctt  | cagcccaacc | 1980 |
| ctgcactgtg | ggagcccctt | tctggggtgg | ccaaggccag | agccggctcc  | ctcagcttgc | 2040 |
| aggggaggtg | ggagggagag | gctcaagcag | gaacccgggc | tgccgacggc  | gottgcccgc | 2100 |
| cagctggagt | tcgggtgggg | cgtgggcttg | gcggccccc  | cactcgggag  | agcgggocag | 2160 |
| ccctgcccag | ccccgggcaa | tgagaggctt | agcaccgggg | ccagcgggtg  | cggaggggtg | 2220 |



|             |             |            |            |             |            |      |
|-------------|-------------|------------|------------|-------------|------------|------|
| caggacagcc  | tggaaacttc  | gatgggtgco | atcaaagtgt | gggggtgggca | caggagccaz | 6000 |
| gaacccaatgt | cottatotca  | ggtaggggot | caggaggtot | cccagacagg  | cagcctccgg | 6060 |
| agagtttggg  | ggtagggaatg | ggagcaacaa | ggcttctttt | ttctctctct  | agaatttggg | 6120 |
| ggotttgggg  | acaggcttga  | gaatcccasa | ggagaggggg | aaagggaact  | ccccacaaag | 6180 |
| tetgcccagag | cgagagaggg  | agaacccgac | toagotgcca | cttccccaca  | ggcct      | 6235 |

<210> 7  
 <211> 278  
 <212> DNA  
 <213> Homo sapiens

|             |            |            |            |            |            |     |
|-------------|------------|------------|------------|------------|------------|-----|
| <400> 7     |            |            |            |            |            |     |
| aagotttttat | aggtgtaaat | tttccactta | gtactgcttt | tgtaatgttg | totTTTTatt | 60  |
| ttcatttato  | tcaagatgtt | ttctaatttc | acttgaottc | otttctaatt | tottacotca | 120 |
| tgtagacata  | catttttggc | ootatgcatt | gggatgcaaa | accagactaa | tttaotttgt | 180 |
| acaaaaagaa  | aatgagaaa  | gaatatatt  | tggtcttgtg | agcactatat | ggaaataott | 240 |
| tatatcccat  | ttgttttato | atattoatat | atcccttc   |            |            | 278 |

<210> 8  
 <211> 73  
 <212> DNA  
 <213> Homo sapiens

|             |            |            |            |            |            |    |
|-------------|------------|------------|------------|------------|------------|----|
| <400> 8     |            |            |            |            |            |    |
| cattgggatac | ttcatcaoct | gctgtgatat | tatgaatgtc | tgcctatata | aatattcact | 60 |
| attccataac  | aca        |            |            |            |            | 73 |

<210> 9  
 <211> 3033  
 <212> DNA  
 <213> Homo sapiens

|             |            |             |            |            |            |      |
|-------------|------------|-------------|------------|------------|------------|------|
| <400> 9     |            |             |            |            |            |      |
| actaacataa  | agctgaaggt | gaataaaaaa  | atcagggtta | gcaaaacaaa | tttccatggg | 60   |
| caaataccaa  | ataaaaagta | aataacttta  | agttccocag | aaaatctgaa | ttgaacgtag | 120  |
| acaaaatgct  | catttctcag | tgcttgacag  | acttaacagt | ttgagccaat | aaaaatgtac | 180  |
| tgactagata  | aactaotaaa | agttgttaat  | ttttgcaatg | tatatctctg | aaaagaaagt | 240  |
| ttatctatata | tagaaattcc | tggtcccaatt | taagaacttt | gagcatttta | attgtttaat | 300  |
| aatatagtct  | aatgtcatca | tgaaataaat  | caataatcaa | atttatttgg | tttattttaa | 360  |
| aaaactgatt  | otttotgctc | tctotatat   | tagactgatt | ttatactaat | gttgcctaaa | 420  |
| gatacaccaaa | ttgtttgaag | ootagggttt  | tgagggatgg | aaaatgatgt | caaaactatt | 480  |
| tacagttcacc | acacacattc | tggggattta  | atacatcctt | tacaagtcca | ggaaagggtg | 540  |
| aagatttgatg | attttggggg | atttagagcta | ccacacccca | gagggtggta | tggatatgtg | 600  |
| tctgttgtga  | gctgtgtgaa | tcagagagtt  | tgattttagc | atatattttg | aaagaggaaa | 660  |
| gatgaaccaa  | tcaaaaataa | taactataat  | gaottttcaa | gatatagaca | atacagttas | 720  |
| gatataaatg  | gaacacaaaa | aagttaaaag  | tggggagatg | aagtctgatt | ttttggtttt | 780  |
| tttttttttt  | tgcttttttg | tttgtttatg  | taatcagtg  | taccagttta | aaataatggg | 840  |
| ttataagaca  | ctatatgcaa | gootcatggt  | aacctccnat | ctaaaacata | caacaaatac | 900  |
| acacaaaata  | aaaaggagaa | attaaaaonc  | accaccagag | aaaatcaoct | acattaaaag | 960  |
| aaagacaaat  | aggasgaaan | taagaagag   | aaggocatca | aataatcaga | aaatgaataa | 1020 |
| caaaatgaca  | ggaataagtc | otcataaata  | ataacattga | atgtaaatgg | actaangctc | 1080 |
| ccaatgaaag  | acagggagtg | gotgaatgta  | ttttaaaaaa | aatatttaac | cgagctgtg  | 1140 |
| gtggtgtctc  | acacctataa | tcccagcatt  | ttgggagact | gagccgggtg | gatcaottga | 1200 |
| gcccaggagt  | tcgagaccag | cctggccaac  | atggcaaaac | cctgtctota | ctaaaaatac | 1260 |
| aaaaaattag  | otgaacatgg | tggcacatgc  | otgtggttcc | agctactaga | gaggctgagg | 1320 |
| cagaagaatt  | gcttgaactt | gggagggtga  | ggttgoagtg | agctaagatt | gatggagcca | 1380 |
| otgaacccca  | gcctaggtga | cagaataaga  | ctctgcctca | aaaaaaaana | gcaaaaacaa | 1440 |
| acaaaacaaa  | aaacccctag | acccaatgat  | toattgccta | caagaagtat | gottcaoctt | 1500 |
| taaaagacaa  | tatagactga | aggtaaaagg  | atggaaaaat | attotatgcc | tatggaaaca | 1560 |
| aaacaaaaga  | agcagaagot | acatttatat  | cagacaaaat | agaotgcaag | acaaaaacta | 1620 |
| tgaaaagaga  | gaaagaaggt | catttatatg  | tgataaagg  | gtccatttag | caagagcatt | 1680 |
| taacaattct  | aaatatatat | tcacccaata  | ctggagtaot | caggtatata | aagcaaatat | 1740 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| aotgggtgcc  | ccagcagtgcc | cagcccgccg  | ggcgtgtgct  | ogctcgattt  | otcaotgggc  | 2280 |
| otttagcagcc | ttcccgcggg  | gcaggggctcg | ggacctgag   | cccgccatgc  | ctgagccccc  | 2340 |
| ootccatggg  | ctcctgtggg  | gcccagagcct | ccccgacgag  | ccccaacccc  | tgctccacag  | 2400 |
| ogccagtcoc  | oatccagccac | gcaagggctcg | agaagtgcgg  | gcgcaacgga  | ccgggaactgg | 2460 |
| caggcagcta  | cccctgcagc  | cotgggtgcgg | aatccactgg  | gtgaagccag  | ctggggtccct | 2520 |
| gagtcctggg  | gagacttgga  | gaacottttat | gtctagotca  | gggatcgtaa  | atacaccat   | 2580 |
| cagcaacotg  | tgtctagctc  | agggtctgtg  | aatgcaacaa  | tccacactct  | gtatctagct  | 2640 |
| actctgatgg  | ggccttgagg  | aacottttatg | tctagotcag  | ggattgtaaa  | tacacccaatc | 2700 |
| ggcactctgt  | atctagotca  | aggttttgtaa | acacacccat  | cagcaacotg  | tgtctagctc  | 2760 |
| agggtatgtg  | aatgcaccaa  | tgcagagctct | gtatctggct  | actttccatg  | gcctccgtgt  | 2820 |
| gaagagacca  | ccaaaacagg  | tttgtgtgag  | caataaagct  | tctatccact  | gggtgcagggt | 2880 |
| gggctgagtc  | cgaaaagaga  | gtcagcgaa   | ggagataagg  | gtggggccgt  | tttataggat  | 2940 |
| ttgggtatgt  | aaagggaat   | tacagtcacaa | gggggtttgt  | tctctggcgg  | gcaggagtg   | 3000 |
| ggggctgcga  | gggtctcagt  | gggggtgctt  | tttgagccag  | gatgagccag  | gaaaaggact  | 3060 |
| ttccaaagg   | aatgtcatca  | attaaaggcaa | ggaccgcgca  | tttacaacct  | ttttgtgggtg | 3120 |
| gaatgtcatc  | agttaagttg  | gggcagggcga | tattcaacttc | ttttgtgatt  | ottcagttac  | 3180 |
| ttcaggccat  | ctgggcgttat | atgtgcgaagt | tacaggggat  | gcgatggctt  | ggcttggggt  | 3240 |
| cagaggcttg  | acagctactc  | tggtggggcc  | tttgagaatg  | tttgtgtoga  | caotctgtat  | 3300 |
| ctagttaatc  | tagtggggag  | gtgggagaa   | tttgtgtcta  | gctcagggat  | tgtaaacgca  | 3360 |
| ccaatcagcg  | ccctgtcaaa  | acagacccact | oggtctatac  | aatcagccag  | atgtgggtgg  | 3420 |
| ggcgagataa  | gagaaataaa  | gcagggctgac | cgagccagca  | gtggcaacgc  | gcacaggttc  | 3480 |
| ctatccacaa  | tatggcagct  | ttgttctttt  | gtgtgttgct  | ataaatcttg  | ctactgctcg  | 3540 |
| otttttgggt  | ccacactgct  | tttatgagct  | gtaacactca  | ccaggaaggt  | ctgcagcttc  | 3600 |
| actcctgaag  | ccactaagac  | ccagagccca  | ccgggagga   | tgaacaaact  | oggcgcgct   | 3660 |
| gccttaagag  | gtataacact  | caocgcgaa   | gtctgcagct  | taactctcca  | gcagccgaga  | 3720 |
| ccaaggaacc  | accagaagg   | agaaactgct  | aacacatctg  | aacatcagaa  | ggaaacaaact | 3780 |
| ccagatgcac  | caccttaaga  | gctgtaacac  | taactgcgag  | ggtccgcggc  | ttcctctctg  | 3840 |
| aagtctagtg  | gaccaagcac  | tcaccagttt  | cggaacacag  | ccagggagtt  | tgagatcagc  | 3900 |
| ctgggcaaca  | tgatgaactg  | cctctctgct  | aaaaaataaa  | aaaattacaa  | aaattggcgg  | 3960 |
| agoatgggtg  | tcctgtcctg  | tggtcccagc  | taogcgggag  | gctaaagtgg  | gaggatcgct  | 4020 |
| tgagcctggg  | agggtgaagc  | tgcaagtggc  | tggtgattgt  | ccacagccct  | ctaggctggg  | 4080 |
| ggcagagctg  | agacccctgt  | tcacccctgc  | aaaaaaattg  | acaaaagtgt  | aataagaggt  | 4140 |
| gcctgatatg  | gctaggcgca  | gtggctcatg  | cctgtaatcc  | cagcaacttg  | ggagcccgag  | 4200 |
| gcgggggggt  | cacctgaagt  | caggagtggt  | agaccagcct  | ggccaaactg  | gagaaagccc  | 4260 |
| atctctctta  | aaaatacaaa  | attagccggc  | tggtggggca  | gtgggtggag  | atgctgttaa  | 4320 |
| ttccagctac  | tcagggaagct | gaggcaggag  | aatcacttga  | acccaggagg  | oggcggttgc  | 4380 |
| agtgagccga  | gatcgtgcc   | ttgcactcca  | ccactccag   | cctgggcaac  | aagagccaaa  | 4440 |
| ctctgtctta  | aaaaaaaaaa  | aaaaaaagtc  | ctgacatata  | agaggtgtgc  | aatgcaatag  | 4500 |
| ttgcagggca  | acatgtttaa  | gaatgtggag  | ctcctgcctt  | ccatggctct  | gttaaaaaac  | 4560 |
| caaccccaag  | gcaggttgca  | gtggctcatg  | ootataatcc  | cagcaacttg  | ggagcccgag  | 4620 |
| gcgggtggat  | cacctgaagt  | caggagttcg  | agaccagcct  | gaacaccaac  | atggtgaaat  | 4680 |
| cccaccccta  | ctaaaaatac  | aaaattagat  | gagcatgggt  | gtgatgctct  | gtaatccac   | 4740 |
| ctacttgga   | ggctgaggca  | tgaaaatcac  | tagaacccag  | gaggcggagg  | ttgtagtgg   | 4800 |
| ccgagatcgt  | gccattgcac  | tcacgcctga  | gcaatgagcg  | aaactccatc  | tcacacacac  | 4860 |
| aaacacaaaa  | acccactctc  | tactccacag  | gagctgggta  | cagagctggg  | ccacatcagt  | 4920 |
| gcaaggtgct  | gagccacaga  | gctaaggcgg  | agctgcaagg  | ccgcggacca  | gataacagt   | 4980 |
| tgtgagatca  | gtgtgtgaga  | tcagagctcc  | ctgcacttgg  | tgacacccag  | ggggccccc   | 5040 |
| agcaacagag  | atggccccc   | ccagtcaccc  | catccacttc  | tcacacagag  | atgtctgttt  | 5100 |
| ottgggaacg  | tggggtaaat  | taggaacaga  | ggtgacagtc  | ttgggtgtgg  | tcagtccagc  | 5160 |
| tgcccccagg  | agcccttgtg  | gcctgtagaa  | aacgttcagg  | cctaggccgg  | gcacggtggc  | 5220 |
| tcacgcctgt  | aatcccaagc  | ctttgggagg  | ccgaggcggg  | tggtcacgga  | ggtcaggaga  | 5280 |
| tcgtgaccat  | cctggctaac  | acggtgaaac  | cccgtctcta  | ctaaaaatac  | aaaaaattgg  | 5340 |
| oogggcatgg  | tgccgggcac  | ctgtagtctc  | agctactcgg  | gaggctgagg  | caggagaatg  | 5400 |
| gcgtgaaccc  | gagaggcaga  | ttttgcaagt  | agccagagac  | gcgcacactg  | actccagct   | 5460 |
| ggggtgacaga | gcaagactcc  | atctggaaaa  | gaaaaagaaa  | acgttcagg   | ctgagccaga  | 5520 |
| ggccccaggct | gtaattctgt  | caottacact  | gacottgggc  | aaggcaacttc | ottccctggc  | 5580 |
| ccagttcacg  | gggttggaat  | cgactccaa   | gtccctccca  | gcattaacgc  | tgcattggtc  | 5640 |
| taagatgaga  | agatggggca  | gtttccctcc  | tctcacccca  | gcocgtgtcc  | acttcaaggt  | 5700 |
| gaatgacacg  | ggaagtcaog  | tgtcccaatc  | ccgaagttcc  | aaagcccttg  | gggacccctac | 5760 |
| tgtcaggggt  | gtgcacagag  | aggtgaaggt  | cagggtgagc  | aatcgccctg  | aagggtcttg  | 5820 |
| cctcattcgg  | gcagagacac  | cgtttctctc  | tggtcttaac  | gggttcttag  | gggttttagc  | 5880 |
| cgaatgagtc  | atggggggcg  | gggggggttc  | tgggggagtt  | ccagctaat   | caacttggga  | 5940 |



|            |             |             |             |             |             |      |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| tattagagoc | aaagagagag  | atagacagao  | ccccatacaa  | taataactgg  | agaotttoac  | 1800 |
| accccacttt | cagcattgga  | cagatoatco  | agacagamaa  | ttaacaaaca  | tcaaatattca | 1860 |
| totgoaacat | aggtcaaatg  | gacotagtag  | atattttacag | aacatttgat  | caaacagotg  | 1920 |
| tagaatacac | attctctotco | toagoacatg  | gataattoto  | aaggetatac  | caaatgctag  | 1980 |
| gtcacaaaao | aaatctttaa  | atttagaaaa  | aaagtgaat   | aatatcaaac  | gtttttctctc | 2040 |
| accacagact | aagaaaaaaa  | gaagtcocaa  | atnaatacaa  | totgagataa  | aaaaggagao  | 2100 |
| gagacaacca | ataccacaaa  | aaattaaagg  | atoattagaa  | gatactatga  | aaotatatgo  | 2160 |
| taataaattg | gaaaaactga  | acaaaastaga | taattcootag | aaacatacaa  | catactggto  | 2220 |
| tgttoagggt | ttgtattttt  | tcatagtacc  | atgaagaant  | acaagaattg  | tttctagaa   | 2280 |
| cattcttcta | ttctcttoatg | gtttttgtat  | ttcttcatgg  | aacctgaag   | aaatacaaaa  | 2340 |
| tgtgaacagg | coaataacaa  | gtaatgagao  | agaagcoata  | ctaaaaagta  | tccacagaaa  | 2400 |
| gaactcagga | tctgatggot  | tcactgatga  | atatttgcaa  | atattttaaa  | aaotaatacc  | 2460 |
| aatcvaactc | aaattattaa  | aaaaatagag  | gtggacagaa  | tottttcaaa  | tgtattctat  | 2520 |
| gaggoaagt  | ttttttctga  | ttgaatctcc  | catttatattt | taatacaca   | taaaaccaga  | 2580 |
| gaaagacaac | ttaaaaagaa  | agaaaactgt  | aggocaatat  | ctctgatgaa  | cattgatgoa  | 2640 |
| gaaatcctca | acaacaaatt  | agcaaaotga  | attoaagaa   | acattaaac   | aatcattcat  | 2700 |
| catgacaaag | tggaaattgt  | cctagagatt  | caagtgtggt  | taggtatgtg  | cagatcaatg  | 2760 |
| ggtttaattg | tgtccaatga  | acataatgtc  | ctcoagctcc  | atocatgtto  | ttgcaaatga  | 2820 |
| caggatctca | ttctttttta  | tggotaagta  | gtactocatt  | gtgtataagt  | gocataatttt | 2880 |
| ctttatccat | tcatctgtta  | gacacotaag  | ttgottccaa  | atotttagota | ttgtgaatag  | 2940 |
| tgotgoata  | aacatgggag  | tgtaaatatt  | ttgttgacat  | aotgatttca  | tttcttttgg  | 3000 |
| ataaatacco | agtagtggga  | ttgctggatc  | ata         |             |             | 3033 |

---

[Translation done.]



## CLAIMS

## [Claim(s)]

[Claim 1] Said complex which is complex which promotes the alteration of the target sequence in a cell, and contains :double-stranded-DNA array, a homologous recombination improver, and \*\* that checks non-homologous edge connection.

[Claim 2] Complex of claim 1 with which a homologous recombination improver is chosen from :Rad52 protein or its functioning fragment, Rad51 protein or its functioning fragment, Rad54 protein or its functioning fragment, and the group which it becomes from those combination.

[Claim 3] Complex of claim 1 whose homologous recombination improver is Rad52 protein or its functioning fragment.

[Claim 4] Complex of claim 1 chosen from the group which \*\* which checks non-homologous edge connection becomes from \*\* which inactivates hMre11, \*\* which inactivates hRad50, \*\* which inactivates Nbs1, \*\* which inactivates hLig4, \*\* which inactivates hXrcc4, and \*\* which inactivates Ku.

[Claim 5] Complex of claim 1 whose \*\* which checks non-homologous edge connection is a Ku deactivator.

[Claim 6] Complex of claim 5 chosen from the group which Ku deactivator becomes from :anti-Ku antibody, Ku joint oligomer, and Ku joint polypeptide.

[Claim 7] Complex of claim 5 whose Ku deactivator is anti-Ku antibody.

[Claim 8] Complex of claim 1 with which a DNA array includes a straight chain DNA array.

[Claim 9] Complex of claim 1 with which at least one targeting array adjoins a DNA array.

[Claim 10] Complex of claim 1 with which a DNA array includes an exogenous control array.

[Claim 11] Complex of claim 10 whose control array is a promotor, an enhancer, an upper activation array, a scaffold adhesion field, or a transcription factor binding site.

[Claim 12] Complex of claim 11 whose control arrays are a promotor and an enhancer.

[Claim 13] Complex of claim 11 whose control arrays are a promotor and an upper activation array.

[Claim 14] Complex of claim 3 with which coating of Rad52 protein or its functioning fragment is carried out on the DNA array.

[Claim 15] Complex of claim 3 Rad52 protein or its fragment of whose is Homo sapiens Rad52.

[Claim 16] Complex of claim 7 whose anti-Ku antibody is anti-Ku70 antibody.

[Claim 17] Complex of claim 7 whose anti-Ku antibody is anti-Ku80 antibody.

[Claim 18] Complex of claim 7 in which at least one anti-Ku antibody is carrying out covalent bond to the DNA array.

[Claim 19] Complex of claim 7 in which at least one anti-Ku antibody is carrying out covalent bond to \*\* which increases homologous recombination.

[Claim 20] Complex of claim 7 with which complex contains anti-Ku70 antibody and anti-Ku80 antibody.

[Claim 21] Complex of claim 9 with which a targeting array originates [ a DNA array ] in the array of 5' of an FSHbeta gene protein coding region, including a control array.

[Claim 22] Complex of claim 9 with which a targeting array originates [ a DNA array ] in the array of 5' of an IFNalpha gene protein coding region, including a control array.

[Claim 23] Complex of claim 9 with which a targeting array originates [ a DNA array ] in the array of 5' of a GCSF gene protein coding region, including a control array.

[Claim 24] Complex of claim 1 which contains further \*\* which checks a mismatch repair protein.

[Claim 25] In the target DNA of a cell, it is the approach of promoting an alteration in the part chosen. : Said approach of introducing into a cell \*\* which increases a double-stranded-DNA array and homologous recombination, and \*\* which checks non-homologous edge connection, promoting the alteration of Chromosome DNA by it, and including promoting an alteration in the part chosen by it in Chromosome DNA.

[Claim 26] The approach of claim 25 that a DNA array includes a straight chain DNA array.



- [Claim 27] The approach of claim 25 that at least one targeting array adjoins a DNA array.
- [Claim 28] The approach of claim 25 that a DNA array includes an exogenous control array.
- [Claim 29] The approach of claim 28 chosen from the group which a control array becomes from :promotor, an enhancer, an upper activation array, a scaffold adhesion field, and a transcription factor binding site.
- [Claim 30] The approach of claim 28 that control arrays are a promotor and an enhancer.
- [Claim 31] The approach of claim 28 that control arrays are a promotor and an upper activation array.
- [Claim 32] The approach of claim 25 that \*\* which increases homologous recombination is chosen from :Rad52 protein or its functioning fragment, Rad51 protein or its functioning fragment, Rad54 protein or its functioning fragment, and the group which it becomes from those combination.
- [Claim 33] The approach of claim 25 that \*\* which increases homologous recombination is Rad52 protein or its functioning fragment.
- [Claim 34] The approach of claim 33 that coating of Rad52 protein or its functioning fragment is carried out on the DNA array.
- [Claim 35] The approach of claim 33 that Rad52 protein or its functioning fragment is Homo sapiens Rad52.
- [Claim 36] The approach of claim 25 chosen from the group which \*\* which checks non-homologous edge connection becomes from \*\* which inactivates hMre11, \*\* which inactivates hRad50, \*\* which inactivates Nbs1, \*\* which inactivates hLig4, \*\* which inactivates hXrcc4, and \*\* which inactivates Ku.
- [Claim 37] The approach of claim 25 that \*\* which checks non-homologous edge connection is a Ku deactivator.
- [Claim 38] The approach of claim 37 that \*\* which inactivate Ku are anti-Ku antibody, Ku joint oligomer, and Ku joint polypeptide.
- [Claim 39] The approach of claim 37 that \*\* which inactivates Ku is Ku antisense molecule.
- [Claim 40] The approach of claim 37 that \*\* which inactivates Ku is anti-Ku antibody.
- [Claim 41] The approach of claim 40 that anti-Ku antibody is anti-Ku70 antibody.
- [Claim 42] The approach of claim 40 that anti-Ku antibody is anti-Ku80 antibody.
- [Claim 43] The approach of claim 40 in which at least one anti-Ku antibody is carrying out covalent bond to the DNA array.
- [Claim 44] The approach of claim 40 in which at least one anti-Ku antibody is carrying out covalent bond to Rad52 protein or its fragment.
- [Claim 45] The approach of claim 25 that a cell is the thing of a fungus, vegetation, or the animal origin.
- [Claim 46] The approach of claim 45 that a cell is the thing of the vertebrate origin.
- [Claim 47] The approach of claim 46 that a cell is the founder or a secondary mammalian cell.
- [Claim 48] The approach of claim 46 that a cell is the founder or a secondary human cell.
- [Claim 49] The approach of claim 46 that a cell is an immortalization mammalian cell.
- [Claim 50] The approach of claim 46 that a cell is an immortalization human cell.
- [Claim 51] The approach of claim 25 which introduces into a cell as complex \*\* which checks \*\* and non-homologous edge connection which increase a DNA array and homologous recombination.
- [Claim 52] The approach of claim 25 which includes further introducing \*\* which checks a mismatch repair protein.
- [Claim 53] A mismatch repair protein: The approach of claim 52 chosen from the group which consists of Msh2, Msh6, Msh3, Mlh1, and PMS2.
- [Claim 54] The approach of claim 52 which is \*\* from which \*\* prevents the manifestation of a mismatch repair protein.
- [Claim 55] The approach of claim 54 that \*\* is an anti-mismatch repair protein antibody.
- [Claim 56] The approach of claim 54 that covalent bond of at least one anti-mismatch repair protein antibody is carried out to the DNA array.
- [Claim 57] The approach of claim 55 that covalent bond of at least one anti-mismatch repair protein



antibody is carried out to Rad52 protein or its fragment.

[Claim 58] The approach of claim 27 that a targeting array originates [ a DNA array ] in the field of 5' of an FSHbeta coding region, including a control array.

[Claim 59] The approach of claim 27 that a targeting array originates [ a DNA array ] in the field of 5' of an IFNalpha coding region, including a control array.

[Claim 60] The approach of claim 27 that a targeting array originates [ a DNA array ] in the field of 5' of a GCSF coding region, including a control array.

[Claim 61] The approach containing the mutation in which Target DNA has less than ten different base pair from a wild type array of claim 25.

[Claim 62] The approach of claim 61 that mutation is point mutation.

[Claim 63] The approach including a wild type array with a DNA array able to correct mutation of claim 62.

[Claim 64] The approach of claim 63 that Target DNA is a cystic-fibrosis film penetration controlling factor (CFTR) gene.

[Claim 65] The approach of claim 64 that mutation changes the amino acid by which a code is carried out to the codon 508 of a CFTR gene coding region.

[Claim 66] The approach of claim 63 that Target DNA is a beta globin gene.

[Claim 67] The approach of claim 66 that mutation changes the amino acid by which a code is carried out to the codon 6 of a beta globin gene coding region.

[Claim 68] The approach of claim 63 that Target DNA is a factor VIII gene.

[Claim 69] The approach of claim 68 that mutation changes the amino acid by which a code is carried out to the codon 2209 of a factor VIII gene coding region.

[Claim 70] The approach of claim 68 that mutation changes the amino acid by which a code is carried out to the codon 2229 of a factor VIII gene coding region.

[Claim 71] The approach of claim 63 that Target DNA is a factor IX gene.

[Claim 72] The approach of claim 63 that Target DNA is a von Willebrand factor gene.

[Claim 73] The approach of claim 63 that Target DNA is a xeroderma pigmentosum group G gene.

[Claim 74] The homologous recombination cell created by the approach of claim 25.

[Claim 75] In a cell, it is the approach of changing the manifestation of the protein coding sequence of a gene. : A DNA array introduces the complex of claim 1 including a control array into a cell.; the bottom of the conditions which enable the alteration of the genome array made into a target, and produce a homologous recombination cell -- a cell -- maintaining --; -- and -- Said approach of enabling the manifestation of the protein coding sequence of a gene under accommodation of a control array, and including maintaining a homologous recombination cell under the conditions which change the manifestation of the protein coding sequence of a gene by that cause.

---

[Translation done.]